BEDINGUNGEN KOOPERATIVEN VERHALTENS: EINE META-ANALYSE

I n a u g u r a l - D i s s e r t a t i o n

zur Erlangung des Doktorgrades

der Philosophischen Fakultät

der Universität Düsseldorf

vorgelegt von

Joachim Tries

aus Oberhausen

1984

BEDINGUNGEN KOOPERATIVEN VERHALTENS: EINE META-ANALYSE

I n a u g u r a l - D i s s e r t a t i o n

zur Erlangung des Doktorgrades

der Philosophischen Fakultät

der Universität Düsseldorf

vorgelegt von

Joachim Tries

aus Oberhausen

1984

D 61

Gedruckt mit der Genehmigung der Philosophischen Fakultät
der Universität Düsseldorf

Referent: Prof. Dr. Lutz F. HORNKE
Korreferent: Prof. Dr. Christine SCHWARZER
Tag(e) der mündlichen Prüfung:
02.11. 1984 Erziehungswissenschaft (Prof. Dr. HORN
06.11. 1984 Soziologie (Prof. Dr. LOWINSKI)
14.11. 1984 Psychologie (Prof. Dr. MANZ)

INHALTSÜBERSICHT

EINLEITUNG

TEIL I THEORETISCHER TEIL

1. Zur Methodik der Forschungsintegration 5

1.1. Methodologischer Exkurs 5

1.2 Methodische Basisannahmen 7
1.2.1. Das Signifikanzkonzept 8
1.2.2. Das Replikationskonzept 10
1.2.3. Verzerrung durch die Publikationspolitik 11

1.3. Methoden der Datensynthese 15
1.3.1. Der Qualitative Übersichtsartikel 18
1.3.2. Statistische Integrationsverfahren auf der Basis des Signifikanzkriteriums 20

1.4. Die Meta-Analyse 32
1.4.1. Grundrationale der Meta-Analyse 32
1.4.2. Literaturrecherche 34
1.4.3. Entwicklung von Forschungsfragen und Erstselektion der Individualstudien 37
1.4.4. Analyse der Einzelstudien 38
1.4.5. Berechnung und Schätzung der Effektgröße 42
1.4.6. Berechnung mittlerer Effekte und Varianzaufklärung 51
1.4.7. Ausgewählte Probleme zur Meta-Analyse 52
1.4.8. Übersicht zu durchgeführten Meta-Analysen zwischen 1976 und 1980 55

2. Zur Kooperationsforschung 58

2.1 Zum Gegenstandsbereich 58

2.2 Theorien zur Erklärung kooperativen Verhaltens 67
2.2.1. Lerntheoretische Erklärungen 68
2.2.2. Austauschtheoretische Erklärungen 78
2.2.3. Das Modell von HAKE 83
2.2.4. Die Kooperationstheorie von DEUTSCH 84
2.2.5. Das Gefangenendilemma als Forschungsparadigma 88

TEIL II DURCHFÜHRUNG DER META-ANALYSE

3. Problemstellung 93

4. Vorgehensweise 95

5. Ergebisse 122
5.1. Deskription 122
5.2. Heuristische Varianzerklärungen 163

6. Diskussion der Ergebnisse 185

Literaturverzeichnis 219

Anhang 252

TABELLENÜBERSICHT

01 Beziehung zwischen Publikationstyp und Ergebnis..... 13

02 Übersicht unterschiedlicher Integrationstechniken..... 17

03 Übersicht der Review-Befunde nach HALL (1978)..... 19

04 Gewichtete und ungewichtete z-Werte aus fünf Studien..... 25

05 Datenbeispiel aus FISHER (1948)..... 27

06 Cluster-Matrix aus der Integration von LIGHT & SMITH (1971)..... 29

07 Ausfallquoten zwischen erstem Literaturpool und tatsächlich in die Meta-Analyse integrierter Studien..... 36

08 Gegenüberstellung verschiedener Prüf-Systeme für die kritische Begutachtung von Individualstudien..... 41

09 Interkorrelationen zwischen unterschiedlichen Effekt-Koeffizienten nach KULIK, KULIK & COHEN (1979)..... 45

10 Beispieldaten für extensive und intensive Treatmentwirkung..... 46

11 Häufigkeiten der untersuchten abhängigen Variablen in 57 Meta-Analysen..... 56

12 Untersuchte unabhängige Variablen in 57 Meta-Analysen..... 57

13 Deutschsprachige Zeitschriften zur Kooperationsforschung..... 95

14 Verteilung der Studien nach Publikationsherkunft..... 96

15 Die Items des Kodiersystems für die Meta-Analyse..... 97

16 Übersicht der in die Meta-Analyse eingeflossenen Studien..... 100

17 Theoriekerne der Untersuchungen in der Meta-Analyse..... 103

18 Kategorien der Fragestellung..... 104

19 Kategorien der Variablenkonstrukte..... 105

20 Rangordnung der UV-Realisation: MODELLIERUNG..... 107

21 Rangordnung der UV-Realisation: BEKRÄFTIGUNGSKONTINGENZEN VOR DEM MESSVORGANG..... 109

22 Rangordnung der UV-Realisation: BEKRÄFTIGUNGSKONTINGENZEN ALTERNATIVER REAKTIONEN VOR DEM MESSVORGANG..... 111

23 Rangordnung der UV-Realisation: KONTINGENZEN WÄHREND DES MESSVORGANGS..... 113

24 Rangordnung der EINTAUSCHWERTE (REVENUEN)..... 115

25 Zunehmender Einfluß der INSTRUKTION auf kooperative Wahlen..... 116

26 Klassifikation der Meßinstrumente..... 118

27 Datenbasis für die Effektgrößen..... 119

28 Berechnungsmodalität der Effektgrößen..... 120

29 Quelle des Mittelwertes..... 120

30 Effektgrößen in Abhängigkeit des Publikationszeitraumes..... 125

31 Tau-Korrelationen zwischen Treatment-Bedingung und Publikationsjahr..... 127

32 Orientierung an unterschiedlichen Referenztheorien im Verlauf des Erfassungszeitraumes..... 128

33 Mittlere Effektgrößen in Abhängigkeit des Publikationstyps..... 128

34 Mittlere Effektgrößen in Abhängigkeit der Publikationsquelle auf der Grundlage verschiedener Meta-Analysen nach M.L. SMITH (1980)..... 129

35 Referenztheorien in Abhängigkeit des Publikationstyps..... 129

36 Untersuchte Konstrukte in Abhängigkeit der Publikationsquelle.... 129

37 Tau-Korrelationen zwischen Publikationstyp und anderen Variablen..... 130

38 Effektgrößen in Abhängigkeit der Anzahl der Effekte pro Studie.. 132

39 Differenzierung der Kategorie: KONTINGENTE BEKRÄFTIGUNG..... 133

40 Differenzierung der Kategorie: MODELLIERUNG..... 134

41 Differenzierung der Kategorie: VARIATION DER MATRIXWERTE..... 134

42 Differenzierung der Kategorie: HINWEISE ZUM KO-AKTEUR..... 135

43 Differenzierung der Kategorie: AUFGABENSTRUKTUR..... 135

44 Effektgrößen in Abhängigkeit der Referenztheorie..... 136

45 Effektgrößen in Abhängigkeit der Konstrukte verschiedener Theorien..... 138

46 Effektgrößen in Abhängigkeit der a posteriori transformierten unabhängigen Variablen..... 139

47 Mittelwerte und Ränge zum MODELLERNEN..... 147

48 Mittelwerte und Ränge zur KONTINGENTEN BEKRÄFTIGUNG KOOPERATIVER REAKTIONEN VOR DEM MESSVORGANG..... 148

49 Mittelwerte und Ränge zur KONTINGENTEN BEKRÄFTIGUNG ALTERNATIVER REAKTIONEN VOR DEM MESSVORGANG..... 149

50 Mittelwerte und Ränge zu DISKRIMINATIVEN HINWEISEN VOR DEM MESSVORGANG..... 150

51 Mittelwerte und Ränge zur KONTINGENTEN BEKRÄFTIGUNG WÄHREND DES MESSVORGANGS..... 151

52 Mittelwerte und Ränge zu DISKRIMINATIVEN HINWEISEN WÄHREND DES MESSVORGANGES..... 155

53 Mittelwerte und Ränge in Abhängigkeit der REVENUEN..... 156

54 Spearman-Korrelationen zwischen Matrixwerten und Effektgrößen... 157

55 Spearman-Korrelationen zwischen Zellenwerten und Prozentrate kooperativer Reaktionen im Naturspiel..... 157

56 Spearman-Korrelationen zwischen Zellenindizes und Prozentrate kooperativer Reaktionen im Naturspiel und für alle Daten..... 158

57 Durchschnittlicher Treatmentumfang und Zusammenhang zwischen Treatmentdauer und Effektgröße..... 159

58 Der Einfluß untersuchungsbedingter Komponenten auf die Effektgrößen..... 166

59 Schrittweise Regression der transformierten unabhängigen Variablen..... 169

60 Prädiktion für Wahlverfahren bei Experimenten mit Erwachsenen... 172

61 Prädiktion für Wahlverfahren mit Kindern..... 173

62 Prädiktion für Handlungsexperimente mit Erwachsene..... 173

63 Prädiktion für Handlungsexperimente mit Kindern..... 174

64 MODELLIERUNG als Basistreatment..... 174

65 KONTINGENTE BEKRÄFTIGUNG VOR DEM MESSVORGANG als Basistreatment. 175

66 KONTINGENTE BEKRÄFTIGUNG ALTERNATIVER KONTINGENZEN als Basistreatment..... 175

67 DISKRIMINATIVE HINWEISE VOR DEM MESSVORGANG als Basistreatment.. 177

68 KONTINGENTE BEKRÄFTIGUNG WÄHREND DES MESSVORGANGS als Basistreatment..... 178

69 Regression der Effektgrößen in Referenz des Bekräftigungsparadigmas..... 179

70 Regression in Abhanggkeit der Referenztheorie des Modellernens.. 179

71 Regression auf der Basis austauschtheoretischer Paradigmen..... 180

72 Regression auf der Basis feldtheoretischer Paradigmen..... 181

73 Regression für feldtheoretische Paradigmen unter Berücksichtigung der Matrixindices..... 182

74 Regression für die Variable INSTRUKTION zur Strategie..... 182

75 Regression für die Variable VARIATION DER MATRIXINDIZES..... 183

76 Interkorrelationen zwischen Matrixindices und Effektgröße..... 183

77 Regression für Information zum Mitspieler..... 183

78 Überblick varianzerklärender Variablen-Gruppen..... 186

79 Effektgrößen in Abhängigkeit verschiedener Konstrukte..... 189

80 Ergebnisse aus der Untersuchung von BRAVER & BARNETT (1978)..... 194

81 Prozentwerte kooperativer Wahlen im Experiment von GARNER & DEUTSCH (1974)..... 201

82 Mittlere Wahrscheinlichkeiten für unterschiedliche Zugfolgen (aus MOREHOUSE, 1966, S. 358)..... 207

ÜBERSICHT ZU DEN ABBILDUNGEN

01 Addition von Wahrscheinlichkeiten nach EDGINGTON..... 23

02 Graphische Darstellung der Effektgrößen..... 32

03 Motivation als unterlegte Metrik für die Dropout-Rate in Abhängigkeit der Curricula (in Anlehnung an GLASS, McGAW & SMITH, 1982)..... 49

04 Formen des Sozialverhaltens..... 62

05 Theoriekerne in der Übersicht..... 68

06 Konsequenzmöglichkeiten und Folgen im Verhalten nach den Paradigmen der operanten Konditionierung..... 71

07 Verhaltenskontrolle als Möglichkeit zur Kooperation nach THIBAUT & KELLEY (1953)..... 80

08 Kooperatie und kompetitive Interdependenzstrukturen nach THIBAUT & KELLEY (1953)..... 81

09 Wahrscheinlichkeiten der Zielerreichung von zwei Personen (nach DEUTSCH, 1962)..... 86

10 Graphische Darstellung der Ergebnisse aus der Untersuchung von GUMPERT, DEUTSCH & EPSTEIN..... 90

11 Theoretische Konsistenz einer Studie..... 102

12 Stem-and-Leaf der Effektgrößen..... 123

13 Veröffentlichungsentwicklung im Erfassungszeitraum..... 125

14 Häufigkeiten der Studien in Abhängigkeit der Anzahl ermittelter Effektgrößen..... 131

15 Beziehung zwischen Kategorie der Fragestellung und der Kategorie der Referenztheorie am Beispiel 'kontingenter Bekräftigung'.... 136

16 Effektgrößen für Modellierung unter Berücksichtigung weiterer Einflußquellen..... 140

17 Mittlere Effektgrößen für KBV unter Berücksichtigung weiterer Variablen..... 141

18 Mittlere Effekte für KBA in Abhängigkeit weiterer Variablen.... 142

19 Mittlere Effekte für DHV in Abhängigkeit weiterer Variablen.... 143

20 Mittlere Effekte für KBW in Abhängigkeit weiterer Variablen.... 144

21 Effektgrößen in Abhängigkeit verschiedener Modellierungs-Niveaus..... 147

22 Effektgrößen in Abhangigkeit verschiedener Niveaus kontingenter Bekräftigung vor dem Meßvorgang..... 148

23 Effektgrößen in Abhängigkeit verschiedener Niveaus kontingenter Bekräftigung alternativer Reaktionen vor dem Meßvorgang..... 149

24 Effektgrößen in Abhängigkeit verschiedener Niveaus diskriminativer Hinweise vor dem Meßvorgang..... 150

25 Effektgrößen in Abhängigkeit verschiedener Niveaus kontingenter Bekräftigung während des Meßvorgangs..... 151

26 Effektgrößen in Abhängigkeit der TFT-Strategie..... 152

27 Effektgrößen in Abhängigkeit kooperativer Reaktionen durch den simultanen Ko-Akteur..... 153

28 Effektgrößen in Abhängigkeit von Display..... 154

29 Effektgrößen in Abhängigkeit verschiedener Niveaus diskriminativer Hinweise während des Meßvorgangs..... 155

30 Effektgrößen in Abhängigkeit verschiedener Revenuen..... 156

31 Effektgrößen in Abhängigkeit des Meßinstruments..... 159

32 Effektgrößen in Abhängigkeit des Alters und des Geschlechts.... 160

33 Effektgrößen in Abhängigkeit verschiedener Studienvaliditäten.. 162

34 Deduktion verschiedener Einflußebenen auf die Effektgrößen..... 164

35 Varianzaufklärung in Abhängigkeit der Population und der Meßverfahren..... 171

36 Suppressionseffekte zwischen DHV und Matrixwert T..... 173

37 Effektgrößen in Abhangigkeit kooperativer Reaktionen durch den Simultanden..... 193

38 Wirkung diskriminativer Hinweise unter Berücksichtigung weiterer Bedingungen..... 199

39 Stem-and-Leaf der Effektgrößen für TFT-Strategien..... 208

40 Untersuchungsergebnisse nach McNEEL (1973)..... 209

41 Kooperationseffekte in Abhängigkeit der Strategie nach WILSON (1971)..... 210

GLOSSAR

AM	Arithmetisches Mittel
AM(ES)	Arithmetisches Mittel der Effektgrößen
AM(T)	Arithmetisches Mittel der Treatmentgruppe
AM(C)	Arithmetisches Mittel der Kontrollgruppe
a, alpha	Irrtumswahrscheinlichkeit Alpha
a1	eine kooperative Wahl der Person A im PDG
a2	eine unkooperative Wahl im PDG
AV	Abhängige Variable
AT	Austauschtheorie
(A)R(G)	eine gebende Reaktion der Person A
(A)K(+)	positiver Verstärker für Person A
b1	eine kooperative Wahl der Person B im PDG
b2	eine unkooperative Wahl im PDG
C	kooperative Wahl im PDG
CC	kooperative Episode im PDG
chi	empirisches
CHI	theoretisches
D	nicht-kooperative Wahl im PDG
DD	kompetitive Episode im PDG
d	Effektstärke nach COHEN
DA	Daten
df	Freiheitsgrade
DHW	Diskriminative Hinweise während des Meßvorgangs
DP	Display
e	(FRICKE)
ES	Effektgröße (effect size)
FTD	Kooperationstheorie nach DEUTSCH
gt	größer als
ge	größer gleich
H	Handlungsmessung
Ho	Nullhypothese
H1	Alternativhypothese
KBA	Kontingente Bekräftigung alternativer Reaktionen
KBV	Kontingente Bekräftigung vor dem Meßdurchgang
KBW	Kontingente Bekräftigung während des Meßvorgangs
lt	kleiner als
le	kleiner gleich
MD	Median
MOD	Modellierung
MS	Mean Square
MI	Motivinduktion

NT	Nachtest
OP	Paradigma der operanten Konditionierung (nach SKINNER)
PDG	Prisoner Dilemma Game
P	Punishment (Zelle 4 im PDG)
p	Wahrscheinlichkeit
PROZRG	Prozentwert kooperativer Reaktionen
PROZRG/S	Prozentwert kooperativer Reaktionen des Simultanspielers
R(G)	gebende Reaktion
R(N)	nehmende Reaktion
R(I)	instrumentelle Reaktion
REV	Revenuen
REFT	Referenztheorie
REFK	Referenzkonzept
RE	Reward (im PDG)
R	Replikation
SD	Standardabweichung
SD(C)	Standardabweichung der Kontrollgruppe
SD(p)	gepoolte Standardabweichung
SDV	Diskriminative Hinweise vor dem Meßvorgang
SDW	Diskriminative Hinweise während des Meßvorgangs
SKLT	Sozialkognitive Lerntheorie
S	Ausbeutung (Zelle 3 im PDG)
SQR	Quadratwurzel
SUM	
T	Versuchungswert (Zelle 2)
TFT	Tit for Tat
TR	Vertrauen (Trust)
**n	Exponent n
UV	Unbhängige Variable
UVV	Summe der unabhängigen Variablen vor dem Meßvorgang
W	Wahlmeßverfahren

EINLEITUNG

Menschen - wie auch andere Lebewesen - sind auf Inter-Aktionen angewiesen, um individuelle oder gemeinsame Ziele zu erreichen, Interessen zu verwirklichen oder komplexe Aufgaben zu bewältigen. Häufig sind dabei die Interaktionspartner voneinander abhängig: beide erreichen ihr jeweiliges Ziel nur, wenn die einzelnen Interakteure ihr Ziel erreichen. DEUTSCH (1949) spricht hier von einer kooperativen Interdependenz, der er eine kompetitive gegenüberstellt: die Zielerreichung des einen verhindert, daß der andere sein Ziel auch erreicht.

'Miteinander Umgehen' beinhalten auch dann Konflikte, wenn Menschen in sozialen Beziehungen gemeinsame Ziele zu erreichen beabsichtigen. Zum einen können sich ihre Zwischenziele unterscheiden, zum anderen können mit gemeinsamen Zielen unterschiedliche Vorstellungsinhalte verbunden sein. Dasselbe gilt für die gemeinsame Arbeit. Die jeweiligen Einzelbeiträge mögen aus verschiedenen Gründen variieren, so daß sich der ein oder andere Interaktionspartner nicht entsprechend seines Beitrages gewürdigt sieht.

Solche Konflikte beinhalten die Gefahr, daß eine anfänglich kooperative Interaktion in eine kompetitive oder in individuelles Nebeneinander umschlägt. Nicht selten werden diese Konflikte "anti"-sozial zu lösen versucht: der Konflikt eskaliert und dem vermeintlichen Gewinner steht der reale Verlierer gegenüber. Zur Erklärung anti-sozialer Interaktionen wurden verschiedene Theorien entwickelt (z.B. Frustrations-Aggressionstheorie, Aggression durch Modellierung). Ihnen stehen nur wenige Konzepte, Modelle oder Theorien gegenüber, die eine "pro"-soziale Konfliktlösung über den Ansatz der Datensammlung hinaus zu erklären in der Lage sind.

Kooperation läßt sich aus diesem Blickwinkel als Form prosozialen Verhaltens interpretieren, durch das Konflikte so gelöst werden können, daß die beabsichtigte oder 'sinnvolle' Zielerreichung möglich wird. Vorauszusetzen ist, daß individuelle Lösungen nicht gewünscht, weniger wirksam oder einfach nicht möglich sind. Aus einem anderen Blickwinkel, nämlich dem der Konflikt-, Verhandlungs- und Spieltheorie stellt Kooperation eine 'rationale' Strategie dar, durch die der eigene 'Gewinn' erhöht werden kann. Gewinn ist dabei nicht nur im pekuniären Sinn zu verstehen, sondern schließt psychische Kosten-Nutzen-Betrachtungen ein.

Eine solche Gewinn-Orientierung, die nur optimal durch kooperatives Verhalten realisiert werden kann, birgt einen zusätzlichen Konflikt in sich. Die Bereitschaft zur Kooperation setzt 'Investitionen' voraus. Ein kooperativer Akt ist nämlich als ein Verhalten zu definieren, durch das der Partner seinem Ziel näher kommt oder eine Verstärkung erhält. D.h., die eigene Zielannäherung oder Bekräftigung hängt ebenfalls von der Kooperationsbereitschaft des Partners oder Kontrahenten ab. Diese Interdependenz ist also durch Bekräftigungsaustausch und Reziprozität der 'Investition' gekennzeichnet. Die kurze Beschreibung macht deutlich, daß kooperative Interak-

tionen 'sensible' Systeme darstellen und es nicht verwundert, daß Kooperation so häufig nicht funktioniert.

Nun könnte man meinen, Kooperation sei wider die soziale Natur einer Gesellschaft, die Leistungsdifferenzierung befürwortet. Das diese Sichtweise zu kurz greift, belegen verschiedene Untersuchungen. Im Rahmen der Konflikt- und Verhandlungsforschung kann aufgezeigt werden, daß unterschiedliche Ziele, Interessen und Motive dann zu langfristigen Nutzen aller am Konflikt Beteiligten führen, wenn diese Konflikte durch 'kooperative Strategien' gelöst werden (vgl. LENG & WALKER (1982). Die sozialbiologische Forschung zeigt, daß helfendes Verhalten weniger als ein zweckfreies, allein durch Normen und Ethik induziertes Verhalten betrachtet werden kann, sondern vielmehr als ein Austauschverhältnis. THOMPSON (1980) zeigt in einer umfassenden Studie, daß Hilfsbereitschaft in dem Maße wächst, in dem der aktuell Helfende annehmen kann, daß ihm in einer ähnlichen Situation von dem Hilfeempfänger auch geholfen werden würde. Die Reversibilität von Situation und Rolle wird in ein Verhältnis zu entstehenden Kosten und Nutzen gesetzt. Desweiteren zeigt AXELROD (1980a, 1980b) in einem originellen Experiment, daß kooperative Strategien nicht nur nicht-kooperativen Strategien überlegen sind - d.h. hier, einen höheren Gewinn erzielen - sondern sich auch langfristig gegenüber nicht-kooperativen Strategien durchsetzen können. Für den Bereich der Schule weisen JOHNSON et al. mit Hilfe einer Meta-Analyse endgültig nach, daß kooperative Unterrichtsstrukturen bei komplexen Aufgaben zu höhren Leistungen führen, als nicht-kooperative Interaktionsmöglichkeiten.

Die Befunde zeigen, daß Kooperation eine wirksame und adäquate Interaktionsstrategie sein kann. Wenig systematisches Wissen liegt jedoch darüber vor, durch welche Bedingungen kooperative Verhaltensweisen erzeugt und aufrecht erhalten werden. Es fehlt zwar nicht an Einzelstudien, die versuchen, unterschiedliche Einflußfaktoren auf Kooperation(sbereitschaft) zu extrahieren. Diese Studien unterscheiden sich jedoch hinsichtlich ihrer theoretischen Anbindungen und damit verbunden, hinsichtlich ihrer zu manipulierenden Variablen, operationalen Definitionen und Untersuchungsmethoden. Diesem Forschungszweig wird somit nicht ohne Grund Konzeptionslosigkeit oder Ekklektizismus vorgeworfen (vgl. auch STAUB 1978).

Ergebnisvielfalt ist nun nicht ein besonderes Problem der Kooperationsforschung. Forschungsdisziplinen, die multikausale Modelle oder Theorien zu validieren versuchen, erzeugen artifiziell vielfältige Daten, die aufgrund differenzierter Variablen, Moderationen und unterschiedlicher Hintergrundtheorien Widersprüche in sich bergen. Dieser Tatsache wird seit langem durch die qualitative Zusammenfassung von Untersuchungen Rechnung getragen. Der "feuilletonistische" Übersichtsartikel (KLAUER, 1981) ist eine Informationsquelle, ohne die sich der akkumulative Prozeß der Wissenschaft nicht nachzeichnen ließe. Dieses Verfahren wurde im Verlauf der 70er Jahre zunehmend kritisiert. LIGHT & SMITH (1971), ROSENTHAL (1978), GLASS (1976) und FRICKE (1977) entwickeln alternative quantitative Techniken, die unter dem

Begriff 'Meta-Analyse" intensiv und extensiv diskutiert werden. Dieser Begriff wurde von GLASS (1976) eingeführt und will im Gegensatz zu den Begriffen "Primär-" und "Sekundäranalyse" deutlich machen, daß es sich bei diesem Verfahren um statistische Methoden handelt, die nicht Rohdaten (neu) analysieren wollen, sondern die in den Forschungsberichten genannten Statistiken als "Rohdaten" aufgreifen, um über eine Vielzahl von Untersuchungen Einflußgrößen auf Untersuchungsmerkmale und -variablen erklären zu können. Dabei bildet die experimentelle Forschungsmethode Basis der Integration. Die Wirkung der unabhängigen auf die abhängige Varibale wird als Effektgröße (effect size) berechnet:

$$ES = (AM(T) - AM(C)) / SD(C)$$

ES = Effektgröße, AM(T) = arithmetisches Mittel der Treatmentgruppe, AM(C) = arithmetisches Mittel der Kontrollgruppe, SD(C) = Standardabweichung der Kontrollgruppe.

Diese 'skalenfreien' Effektgrößen können in Abhängigkeit der die Fragestellung(en) zugrundeliegenden Variablen gemittelt werden. Die Varianz der mitleren Effektgrößen ist mit Hilfe multivariater Regressionsverfahren beschreibbar. Dabei fließen nicht nur experimentell kontrollierte Variablen in die Regression ein, sondern auch solche Merkmale, von denen ein Einfluß auf die Effektgröße vermutet wird: Probandenmerkmale, Untersuchungs- und Auswertungsmerkmale, Experimentatormerkmale.

Die Meta-Analyse ist vielfältig einsetzbar. Sowohl ist es möglich, mit ihr die Wirkung implementierter Technologien oder Programme - z.B. Unterschiede zwischen Gesamtschulen und traditionellen Schulen - über viele Evaluationsergebnisse hinweg herauszuarbeiten, wie die Meta-Analyse zur Überprüfung von Hypothesen und Theorien, einschließlich Hintergrundtheorien, angewandt werden kann.

Die vorliegende Arbeit verfolgt im ersten Teil zwei Ziele. Zunächst liefert sie eine Darstellung statistischer Integrationsverfahren unter besonderer Berücksichtigung der Meta-Analyse. Anschließend werden im 2. Kapitel solche Theorien dargestellt, auf die sich empirische Untersuchungen zur Kooperationsforschung in unterschiedlicher Ausprägung und Stringenz berufen. Der zweite Teil beinhaltet die Anwendung der Meta-Analyse auf empirische Untersuchungen zur Koopertionsforschung. Im Gegensatz zur Meta-Analyse von JOHNSON et al. (1981) bildet hier Kooperation die abhängige Variable. Ziel ist es, in enger Orientierung an einem funktional-analytischen Verhaltensmodell die Widersprüche in diesem Forschungsgebiet aufzulösen. Die Analyse der Individualstudien und die Transformation der in ihnen dargestellten Informationen in Variablen auf der Ebene dieses Modells wird besondere Aufmerksamkeit gewidmet.

Die Arbeit entstand in dem Bewußtsein, daß weniger die aktuellen Ergebnisse, sondern vielmehr ihre zukünftige Bewährung oder Widerlegung von Interesse sind.

TEIL I

THEORETISCHER TEIL

1. ZUR METHODIK DER FORSCHUNGSINTEGRATION

1.1 Methodologischer Exkurs

Die Entwicklung und Anerkennung wissenschaftlichen Wissens stellt sich verschiedenen Konventionen, die zugleich einen Einblick in den Prozeß der Wissenschaft liefern. Wissenschaftstheorien - als philosophische Sätze über die Regeln und die Abgrenzung der Wissenschaft - wurden verschieden dezidiert klassifiziert. SPINNER (1977) unterscheidet zwischen 'certistischer Rechtfertigungsrationalität' und 'fallibilistischer Kritikrationalität'. LAKATOS (1974) liefert eine Übersicht über die Entwicklung der empirischen Wissenschaft. Er unterscheidet vereinfacht zwischen Induktionismus, Falsifikationismus sensu POPPER und 'raffiniertem' Falsifikationismus. Während der Indiduktionismus als Rechtfertigungsmethode der Daten- und Evidenzakkumulation betrachtet werden kann, bildet der POPPERsche Falsifikationismus die methodologische Basis des von LAKATOS explizierten Forschungsprogramms. Bei POPPER (1964) fehlen insbesondere Hinweise auf ein Enscheidungskriterium, unter welchen Anomalie- und Datenvoraussetzungen eine Theorie als falsifiziert gilt und unter welchen Bedingungen Elaborationen bzw. Modifikationen im Hypothesennetzwerk möglich sind. KUHN (1969) bot damit nicht ohne Logik und Kenntnis der Wissenschaftsgeschichte eine soziologische Lösung an: Theorien unterliegen revolutionären Zyklen, während dieser sie durch personale Entscheidungen verworfen bzw. durch andere Theorien ausgetauscht werden. LAKATOS vertritt demgegenüber (forschungs-) logische Entscheidungregeln. Sein Konzept von der positiven und negativen Heuristik führt zur Differenzierung zwischen dem als vorläufig 'wahr' akzeptierten Theoriekerns und des zu überprüfenden Hypothesengürtels. Die negative Heuristik beinhaltet die Regel, den modus tolens nicht gegen den Theoriekern zu richten, während die Hypothesen nahezu mit "dadaistischer Phantasie" (vgl. FEYERABEND 1976) und methodischer Akribie überprüft und verändert werden dürfen. Dies bezeichnet LAKATOS als positive Heuristik. Ein solches Forschungsprogramm erlaubt auch und gerade den Umgang mit solchen Wissenschaften, die inzwischen über viele Daten und wenig gültiges Wissen (Theorien) verfügen. Als 'wahr' etikettierte - und dies meint, als noch nicht falsifiziert erklärte - Theoriekerne können herreichen, um die Fülle beobachteter Phänomene und sie erklärender Variablen zu integrieren. Damit sind die Inkommensurabilitäts- und Operationalisierungsprobleme wissenschaftstheoretisch allerdings nicht gelöst.

Einen erhellenden Beitrag zur methodischen Lösung liefert die Meta-Analyse. Ihr zentraler Term, die Effektgröße, d.h., die abgeschätzte Ursachen-Wirkungs-Beziehung bildet die hypostasierten Wirkungen in Beziehung gesetzter Variablen ab. Unterschiede durch verschiedene Meßinstrumente werden adjustiert, so daß über viele Untersuchungen hinweg die Varianz der Effekt-

größe durch begründbare Hintergrundvariablen erkannt und bestimmt werden kann.Aber nicht nur vor diesem Hintergrund liefert dieses Verfahren einen Beitrag zur Lösung wissenschaftstheoretischer Probleme. Sie erlaubt viel grundlegender, Ergebnisse des normalwissenschaftlichen Alltags in ein Forschungsprogramm umzusetzen. Nach BERGMANN (1957) sind sozialwissenschaftliche Theorien weder vollständig noch abgeschlossen. Folglich können Phänomene auch nur im Ausschnitt erklärt werden. Untersuchungen zu demselben Phänomen müßten in ihrem 'Ausschnitt' immer deckungsgleich sein, wenn sie dieselben Daten erzeugen wollten. Das Gegenteil ist in der Forschungspraxis jedoch der Fall: Ergebnisse variieren in Abhängigkeit der spezifischen Untersuchung. Die an den 'harten' Naturwissenschaten orientierten Wissenschaftstheorien liefern zwar die Logik des Daten-Widerspruchs, jedoch noch lange nicht eine Methode, um mit diesem Widerspruch progressiv umgehen zu können. Die Forschungsgemeinschaft steht dann vor einem Entscheidungsdilemma, wenn sie z.B. die Ergebnisse aus einer Vielzahl von Untersuchungen vor dem Hintergrund bewerten soll, daß die Hälfte der Ergebnisse die zu überprüfende Hypothese/Theorie bewährt, die andere Hälfte die Falsifikation der Hypothese/Theorie verlangt. Eine Entscheidung ist nur zu treffen, wenn auf der Ebene der Hypothesen Widersprüche erkannt bzw. erklärt bzw. Mängel in einzelnen Studien nachgewiesen werden können. Die systematische Synthese von Untersuchungsergebnissen liefert eine solche Möglichkeit. Sie erlaubt darüberhinaus die Rationalisierung der Falsifikation. Die Widerlegung einer Theorie orientiert sich nicht allein am Binnen-Wissen über diese Theorie (vgl. KUHN, 1963, LAKATOS, 1974, POPPER,1964), sondern gleichsam am Wissen um die Gültigkeit konkurrierender Theorien. Eine Wissenschaftstheorie wie eine Wissenschaftsmethode müssen daher Rationale bereitstellen, mit deren Hilfe sowohl Theorien (deskriptiv) verglichen und ihre jeweiligen Überlegenheiten (qualititativ und heuristisch-quantitativ) angezeigt werden können.

Insbesondere in den Sozialwissenschaften erklären häufig verschiedene Theorien dieselben Phänomene. Zugleich herrscht Konsens über den unterentwickelten Zustand sozialwisssenschaftlicher Theorien (vgl. IRLE 1975). Demnach wird die negative Heuristik institutionell gehandhabt, ohne jedoch in die Entscheidungsphase mit dem Ziel treten zu können, die unterlegene, d.h. empirisch weniger bewährte wie weniger gehaltvolle Theorie ad acta legen zu können. Wenig Übereinkunft herrscht über die Frage, ob theoretische Außenkriterien (d.h. die bessere Theorie) als Rationalitätskriterium für den Theorievergleich zur Verfügung stehen. HANSON (1965), FEYERABEND (1976) und KUHN (1969) negieren die Möglichkeit, Theorien zu vergleichen: die in den Theorien verwandten Begriffe sind Theorie-verzerrt und folglich unvergeichbar. GIESEN & SCHMIDT (1978) skizzieren diesen Standpunkt:

"Wenn T durch eine andere Theorie T' modifiziert oder ersetzt wird, ändert sich folglich die Bedeutung von S auch dann, wenn S sowohl in T wie in T' Verwendung findet" (S. 234).

Hierin verbirgt sich das SNEED'sche Argument, nach dem theoretische Begriffe empirischer Theorien nur durch den Rückgriff auf das Beweisverfahren dargestellt werden können (vgl. GRECA 1978, S. 272). Theoretische Evidenz liegt damit immer nur auf der Basis der Validität (und das heißt auch wieder aufgrund der theoretischen Evidenz) der Hintergrundtheorien vor. Ein forschungs-logischer Theorienentscheid verlangt demnach Hintergrundtheorien hoher Validität. Ihre Gültigkeit ist aber - zumindest in jungen Disziplinen - ebensowenig eindeutig geklärt wie die Gültigkeit der zu validierenden bzw. zu falsifizierenden Theorien. Es fehlt somit ein forschungs-logisches Entscheidungsrational. Während in 'gefestigten' (Natur-)-Wissenschaften ein Theorieaustausch auf der Basis des erweiterten Informationsgehaltes (T' involviert T) und des höheren Bewährungsgrades (T'erzeugt weniger Anomalien als T) möglich geworden ist, fehlt es den Sozialwissenschaften sowohl an einer methodologischen Konsistenz wie an Forschungsprogrammen, die die systematische Erklärung eines Phänomens erlauben. Die Anerkennung von Theorien ist damit eher an solche Kriterien gebunden, wie sie KUHN (allgemeingültig) erklärt: Theorien leben und sterben mit ihren Protagonisten. Das Modell des Forschungsprogramms im Sinne von LAKATOS begründet jedoch im Gegenteil die negative Heuristik durch die Theoriekonkurrenz: der Theoriekern ist solange nicht aufzugeben, bis eine neue Theorie vorliegt, die (i) all die Phänomene erklären kann, die ihre Vorgängerin auch zu erklären in der Lage war und (ii) darüberhinaus einen höheren Informationsgehalt und Bewährungsgrad besitzt (vgl. LAKATOS 1974). Darin liegt einer der wesentlichen Unterschiede zwischen LAKATOS und POPPER. POPPER (1964)falsifiziert durch ein Innenkriterium, LAKATOS durch ein Außenkriterium. Für die sozialwissenschaftliche Theorienentwicklung ist diese Vorgehensweise dann fruchtbar, wenn es möglich wird, Forschungsdaten zu integrieren. Integrieren heißt dann vielmehr Forschungsergebnisse auf Unterschiede innerhalb wie zwischen den Studien zurückführen zu können, um sie anschließend als Test einer Theorie zu betrachten.

1.2.1 Methodische Basisannahmen

Wie oben skizziert wurde, diskutiert die Wissenschatfstheorie über dezidierte Wissenschaftskonventionen für den laufenden Wissenschaftbetrieb. Mit ihrem meta-theoretischen Gebäude zu leben heißt gleichsam, über solche Kriterien zu verfügen, die Hypothsenentscheidungen erlauben. Gängiges Entscheidungskriterium, mit dem empirische Hypothesen getestet werden, ist das Signifikanzkriterium. Wissentlich, daß eine singuläre Hypothesenentscheidung den Irrtum in der Regel zu 5% in sich trägt, soll das Replikationsparadigma diesen Irrtum minimieren. Beide Entscheidungskonzepte beinhalten Probleme, wenn sie unreflektiert als methodische Konvention gehandhabt werden. Häufig wird der Vorwurf erhoben, daß veröffentlichte (affirmative) Wissen repräsentiere nicht das gesamte (falsifikatorische) Wissen, das in

Form nicht veröffentlichter Papiere, Kongreßunterlagen usw. zur Verfügung stehe. Die Information zur Bewährung von Theorien und/oder Hypothesen sei verzerrt.

1.2.1 Das Signifikanzkonzept

Mit dem Signifikanzkriterium wird die Wahrscheinlichkeit festgelegt, unter der ein unter der Forschungshypothese (H1) formulierter Zusammenhang zufällig (i.S. der Nullhypothese Ho) auftritt. Dieser Irrtum, einen Zusammenhang zu behaupten, wird als Typ I - Fehler bezeichnet. Die Aussage dieser hypothesenentscheidenden Operation gilt für die definierte Population. Allgemein prüft dabei der Signifikanztest, inwieweit eine beobachtete Differenz zwischen zwei zufällig aus einer gemeinsamen Population gezogenen Gruppen nach einer Treatmentapplikation (im Falle des Experiments) durch den Stichprobenfehler bzw. durch den Zufall erklärt werden kann, wenn unterstellt ist, die beiden Gruppen gehörten auch nach der Treatmentapplikation zu derselben Population (Ho). Liegt der berechnete Wahrscheinlichkeitswert unterhalb eines durch Konvention festgelegten minimalen Irrtumskriteriums, ist die Wahrscheinlichkeit entsprechend diese Wertes groß, daß diese Differenz innerhalb derselben Population auftritt. Diese Wahrscheinlichkeit p(D/Ho) gibt keine Auskunft über die Wahrscheinlichkeit, mit der die Forschungshypothese (H1) angenommen werden kann: p(D/H1). Wird die Nullhypothese verworfen, kann induktiv geschlossen werden, daß der determinierende Einfluß des Treatmens (soweit weitere Störfaktoren ausgeschlossen und andere relevante Variablen kontrolliert sind) in der Ursprungs-Population, aus der sowohl die Experimental- wie die Kotrollgruppe gezogen wurden, wirkt und das umgekehrt die Veränderung der Experimental-Stichprobe zu keiner veränderten Zielpopulation führt.

Verschiedene Autoren (BAKAN, 1966, CARVER, 1978, CRONBACH & SNOW, 1977) weisen auf häufig Fehlinterpretationen dieses Tests hin. Als "odd against chance fantasy" bezeichnet CARVER (1978) den Irrtum, das Signifikanzkriterium bezeichne die Wahrscheinlichkeit, unter der eine Hypothese unter gegebenen Daten wahr ist (p(Ho/D). Es wird jedoch nicht danach gefragt, wie groß die Wahrscheinlichkeit einer bestimmten Datendifferenz zwischen zwei Gruppen aus derselben Population ist, sondern wie groß die Wahrscheilichkeit ist, daß zwei Gruppen derselben Population angehören, d.h. mit der Ho unter der Bedingung der Datendifferenz wahr ist. Der "replicability or reliability fantasy" unterliegen ebenso eine Anzahl von Interpretatoren.. Die statistische Wahrscheinlichkeit p = 1 - alpha wird als Wahrscheinlichkeit interpretiert, mit der dieselben Ergebnisse durch Replikation erhalten werden. Hier wird angenommen: p(R/D') = 1 - alpha. Replizierbarkeit hängt jedoch nicht vom Signifikanztest, sondern von der Kontrollierbarkeit und Manipulierbarkeit der Daten ab. Eine letzte Fehlinterpretation nach CARVER liefert die "valid research hypothesis fantasy". Sie unterstellt einen unmittelbaren Zusammenhang zwischen dem Signifikanzkriterium alpha und der

Evidenz für die Forschungshypothese: p (H1/D) = 1 - p (D/Ho).CARVER (1978) verweist auf die Schwierigkeit, mit der diese Fehlinterpretation nachgewiesen werden kann. Die Tradition, häufig p-Werte statt Statistiken zu veröffentlichen, wie das Verständnis über die negative Korrelation zwischen Signifikanzkriterium und Validität der Forschungshypothese (H1) seien jedoch Indikatoren hierfür. BAKAN (1966) schließt, daß mit Hilfe des Signifikanzkriteriums über die Anwendungshypothesen keine Information gewonnen wird. Auch kann über den Bewährungsgrad der theoretischen Hypothesen im Rahmen der Einzelstudien keine Aussagen gemacht werden.

Das Signifikanzkriterium als Entscheidungsrational bereitet jedoch nicht nur auf der Interpretationsebene Mißverständnisse. Der Signifikanztest wird durch fünf Faktoren beeinflußt, von denen die Signifikanzentscheidung abhängig ist (vgl. COHEN, 1966, BAKAN, 1966,GREENWALD, 1975).

Einseitige vs. zweiseitige Testentscheidung
Die Irrtumswahrscheinlichkeit entspricht dem Flächenanteil einer Wahrscheinlichkeitsdichte unter der Normalkurve, die die Auftrittswahrscheinlichkeit unter Annahme der Nullhypothese bestimmt. Bei einseitiger Fragestellung liegt diese Fläche rechts ODER links vom geschätzen Populationsmittelwert, während sie bei der zweiseitigen Fragestellung als Alpha/2 rechts UND links liegt. Die Ablehnungswahrscheinlichkeit bleibt konstant. Jedoch ändern sich die Ablehnungsbereiche. So ist beim zweiseitigen Test eine extreme Abweichung eher mit Ho vereinbar als beim einseitigen Test. Damit ist der zweiseitige Test bei gleicher Irrtumswahrscheinlichkeit strenger.

Signifikanzniveau
Die Signifikanzentscheidung liegt innerhalb der Konventionen im Ermessen des Forschers. Eine stillschweigende Übereinkunft besagt demnach, das 5%-Niveau sei ein gutes, das 1%-Niveau sei ein besseres Kriterium (vgl. GREENWALD, 1975). Durch die Computerauswertung wird die empirische Signifikanz gleich mit berechnet. Diese a posteriori-Inspektion verletzt das Inferenz-Modell.

Umfang der geschätzten gemeinsamen Fehlervarianz
BAKAN bezeichnet die Fehlervarianz als eine Gegebenheit der Situation. Homogenisierung der Varianz sind ggf. notwendige Verfahren, solange sie nicht davon ablenken, daß ein Treatment eventuell keinen oder nur geringen Effekt auf den Mittelwert, jedoch eine Veränderung der Varianz zur Folge hat.

Differenz zwischen Experimental- und Kontrollpopulation
Diese Differenz ist die einzige Unbekannte im Ensemble der Einflußfaktoren auf die Signifikanzentscheidung.

Stichprobenumfang
Mit zunehmendem Stichprobenumfang wird die kleinste Differenz statistisch signifikant. Hierin liegt die methodisch einfachste Manipulierbarkeit von Entscheidungen:

"... if the null hypothesis is not rejected, it is usually because the N is too small. If enough data are gathered, the hypothesis will generally be rejected. If rejection of the null hypothesis were the real intention in psychological experiments, these usually would be a need to gather data" (NUNNALLY, 1960, S. 643).

Die Relativierung des Signifikanztests, seine häufigen Fehlinterpretationen sowie die eigentliche Suche nach der Wahrscheinlichkeit von H1 unter gegebenen Daten (p(H1/D)) führen zur Kontroverse über den Stellenwert dieses Tests. TYLER (1931) hebt hervor, daß die statistisch signifikante Differenz nicht notwendig auch eine wichtige Differenz ist, während GOLD (1969) sie als notwendig aber nicht hinreichend für die Wichtigkeit eines Ergebnisses bezeichnet. Demgegnüber weisen WICH & CAMPBELL (1969) die Kritik gegen den Signifikanztest zurück.

"We reason that it is very important to have a formal and nonsubjective way of decision whether a given set of data shows haphazard or systematic variation" (S. 143).

Für die Eliminierung des Signifikanztests plädieren LYKKEN (1968) und CARVER (1978). CARVER argumentiert:

"A statstical significant result erroneously suggests that the probability is low that the result was caused by chance, and that the probability is high that the result is reliable, valid, and important. Even without the fanciful interpretations, statistical significance testing usually involves serious breach of the scientific method; therefore it is more accurate to say that statistical significance testing uses a corrupt form of the scientifc method" (CARVER 1978, S. 396).

1.2.2 Das Replikationskonzept

Während der Signifikanztest Aussagen über die Wahrscheinlichkeit von beobachteten Daten unter der Bedingung der Gültigkeit der Nullhypothese (p(D/Ho)) zuläßt, baut der wissenschaftliche Erkenntnisprozeß auf eine 'intuitive' Wahrscheinlichkeit der Gültigkeit der Forschungshypothese unter der Bedingung der gegebenen Daten (p(H1/D)) auf. Diese Gültigkeit wird durch Repliaktionen in ihrer Intuition erhöht: multiple Tests konfirmieren die Forschungshypothese. Der damit verbundene Replikationsgedanke ist jedoch nicht unumstritten. CARVER (1978) fordert die exakte Wiederholung von Experimenten:

"Whether or not the results are replicable or reliable actually depends upon whether the important variables can be controlled and manipulated in exactly the same way in order to give the same results" (S. 385).

Die exakte Wiederholbakeit oder Wiederherstellbarkeit DERSELBEN relevanten Variablen wird jedoch in Frage gestellt (vgl. N. SMITH (1970, GLASS, 1978).

"Same (kursiv bei GLASS) is not defined, and the respects in which comparable studies must be the same are unspecified. The claim that only studies which are the same in all (kursiv bei GLASS) respects can be compared is selfcontradictory; there is no need to compare them, since they would obviously have the same findings within statistical error" (S. 357).

SIDEMAN (1952) betont stattdessen die Implikationen der Daten zu replizieren:

"Instead of attempting to replicate the data themselves, one may achieve replication be investigating the implications of the data" (S. 135).

SIDEMAN unterscheidet daher zwischen direkter und systematischer Replikation. In der direkten Replikation wird versucht, die wesentlichen Bedingungen experimentell zu wiederholen, während in der systematischen Replikation zusätzliche Daten zur Erklärung des Phänomens hergestellt werden sollen. Er schließt damit an das wissenschaftstheoretische Konzept von BERGMAN an: unvollkommene Theorien implizieren a priori unkontrollierte und/oder unbekannte Variablen, die eine erfolgreiche Wiederholung des Experiments beeinflussen. Insbesondere das von SIDEMAN als 'Replikation durch Annahme der Konsequenz' bezeichnete Verfahren erlaubt i. S. des Forschungsprogramms von LAKATOS unter Beibehaltung der Theorie sowohl ein Suchen nach genügend Falsifikationskraft wie auch ein Suchen nach weiteren hinreichenden Bedingungen.

N. SMITH (1970) zeigt die Faktoren auf, die bei einer Replikation berücksichtigt werden müssen: (i) die Untersuchungspersonen, (ii) die Apparate, (iii) das Untersuchungsmaterial, (iv) das Untersuchungverfahren, (v) die physikalischen Gegebenheiten, (vi) die psychologische Atmosphäre, (vii) der/die Experimentator(en), sowie (viii) die Interaktion zwischen Experimentator(en) und Versuchspersonen. Durch vielfältige Untersuchungen ist nachweisbar, daß diese Fakoren auch tatsächlich Einfluß auf die Ergebisse ausüben können. (N. SMITH 1970, S. 973).

1.2.3 Verzerrung durch die Publikationspolitik

Das Problem der Sicherheit des Wissens über den Gehalt und die Bewährung von Theorien oder Hypothesen wird durch die kontrovers diskutierte Annahme vermehrt, die wissenschaftlichen Publikationsorgane veröffentlichen insbesondere Beiträge, deren Ergebnisse signifikant sind (vgl. STERLING 1959, BOZARTH & ROBERTS 1972, WAY & LARIMOR 1973, LYKKEN 1968,BAKAN 1966, GREEN-

WALD 1975. N. SMITH 1970, ROSENTHAL 1980). Zwei Ausschnitte repräsentieren diese Kontroverse. WAY & LARIMOR (1973) berichten über die Erfahrung, daß "... well written , clear and concise well designed and adequately analyzed..." Artikel mit der Begründung "... that these results did not allow for the rejection of the null hypothesis ..." (S. 361) zurückgewiesen werden. STERLING (1959) zeigt auf, daß von 362 Forschungsberichten 81% Signifikanztests anwandten, von denen 97% die Nullhypthese verwarfen. WAY & LARIMOR (1973) schließen daraus:

"It may well be that 95% of the replications resulted in nonsignificant results...only one or two of the 5% ... will be published" (S. 362).

In der Konsequenz führen diese Aussagen zu der Vermutung, alle Theorien seien bereits auf der Basis nicht veröffentlichter Ergebnisse falsifiziert. Entsprechend wird versucht, das Dunkelziffer-Argument zu widerlegen. HUNT (1975) findet im Gegensatz zu STERLING , daß von 110 veröffentlichten Tests 29% auf dem 5%-Niveau oder höher signifikant waren. M. SMITH (1980) weist in ihrem Überblick zu verschiedenen - auch nicht-publizierte Studien einbeziehenden - Meta-Analyse nach, daß publizierte und nicht-publizierte Ergebnisse nicht unbedingt entgegengesetzte Ergebnisse erzielen müssen. Der mittlere Effekt für publizierte Studien liegt bei ES = .64 und für unpublizierte Studien im Durchschnitt bei ES = .58. Die Verzerrung beträgt demnach (((.64 - .58)/.58)*100) = 10%.

Tabelle 1: Beziehung zwischen Publikationstyp und Ergebnis
(aus: M.L. SMITH, 1980, S. 23)

Investigator(s)	Topic		Journal	Unpub.
KAVALE (1979)	Psycholinguistic training	n:	13	5
		AM(ES)	.50	.37
HARTLEY (1977)	Computer-based instruction	n:	34	34
		AM(ES)	.36	.54
	Tutoring	n:	9	17
		AM(ES)	.77	1.05
SMITH (1980a)	Sex bias in psychotherapy	n:	28	
		AM(ES)	1.08	
SMITH (1980b)	Effects of aesthetics education on basic skills	n:	29	56
		AM(ES)	1.08	.50
CARLBERG (1979)	Special education room placement vs regular room	n:	146	114
		AM(ES)	-.09	-.14
	Resource room placement vs regular room placement	n:	33	
		AM(ES)	.32	
MILLER (1979)	Drug therapy of psychological disorders	n:	336	
		AM(ES)	.49	
HEAROLD (1979)	Effects of TV on anti-social behavior	n:	262	13
		AM(ES)	.40	.23
SUBTOTALS		n:	1025	268
		AM(ES)	.38	.27
.27				
SMITH, GLASS and MILLER (in press)	Psychotherapy	n:	1179	61
		AM(ES)	.87	1.96
TOTALS		n:	2204	329
		AM(ES)	.64	.58

Die in Zeitschriften publizierten Ergebnisse liegen um .1 Standardabweichungen günstiger für publizierte als für nicht publizierte Untersuchungen

ROSENTHAL (1969), ROSENTHAL (1980) berechnet die Anzahl von Studien, die notwendig sind, um die Null-Hypothese anerkennen zu können. Hierzu wird die Anzahl der Studien berechnet, die im Durchschnitt ein Null-Ergebnis ergeben, bevor die über-alles-Wahrscheinlichkeit für einen Alpha-Fehler das Signifikanzniveau erreicht. Demnach gilt für die Anzahl neuer, 'fehlender' Studien (x):

$$Z_{(\alpha/H_0)} \qquad \frac{K\,\bar{Z}_K}{\sqrt{K+X}}$$

$$\bar{Z}_K = \frac{\sum Z}{\sqrt{K}}$$

$$X \;=\; \frac{K^2\,\bar{Z}_K}{Z^2_{(\alpha/H_0)}} \;=\; \frac{K}{Z^2_{(\alpha/H_0)}} * \left[K\,(Z_K)^2 - Z^2_{(\alpha/H_0)}\right]$$

Am Beispiel der Zusammenfassung von 94 Experimenten, die Effekte der interpersonellen sich selbst erfüllenden Prophezeihung überprüften, ermittelte ROSENTHAL (1969) einen mittleren z-Wert von AM(z) = 1.014 (Sum (z) = 1.014 * SQR(94) = 9.83). Für Z (alpha/Ho) = 1.645 sind entsprechend

$$X = \frac{94}{1.645} * \frac{94}{1104} - 1.645 = 3263$$

Studien mit einem AM(z) = .00 notwendig.

BABE (1978, dtsch 1979) legt damit endgültig das Dunkelziffer-Argument zur Seite.

"Damit sehen wir uns einer prinzipiellen Ungewißheit gegenüber, die alle Bemühungen, in den Sozialwissenschaften gesicherte Daten zusammenzutragen und sie zu interpretieren, zur Augenwischerei degradiert" (S. 21).

1.3 Methoden der Datensynthese

Der wissenschaftsmethodische Konsens verlangt die multiple Testung von Hypothesen mittels systematischer Replikation. Ziel dieses Pozesses ist es, den theoretischen Fortschritt durch Evidenz-Akkumulation in der Absicht fortzuschreiben, größeres Vertrauen in die Generalisierbarkeit von Theorien bzw. Hypothesen zu erlangen wie auch Anomalien zu erkennen und festschreiben zu können, die die Falsifikation einer Theorie einleiten. Sozialwissenschaftliche Theorien generieren 'weiche Daten' durch (1) Ausschluß eventuell interagierender Parameter sowie durch die (2) Anwendung nicht weniger weicher Hintergrundtheorien. WALBERG & HAERTEL (1980) ziehen daraus den Schluß,

"in educational research and evaluation , for example, the number of possible causes and effects seems to have multiplied; and recent claims that instructional and psychological treatments are narrowly effective only under certain conditions, at certain times, and for particular cases of people have encouraged the false equating of complex interactions with scientific process" (S. 5).

Dieser wissenschaftlichen Alltagssituation versuchen Integrationstechniken mit unterschiedlichen Implikationen und Zielen zu begegnen. FELDMAN (1971), JACKSON (1980), und PILLEMER & LIGHT (1980) listen mögliche Ziele auf, die mit solchen Integrationstechniken eingelöst werden können:

(1) die Schaffung eines Kompendiums über einen Forschungsbereich durch die Aufzeichnung inhaltlicher und methodischer Entwicklungen auf diesem Gebiet

(2) Überprüfung von Theorien bzw. Hypothesen durch die Synthese der Einzelinformationen zu Wissen

(3) Beschreibung der Art von Beziehungen in Abhängigkeit unterschiedlicher Moderatorvariablen

(4) Entwicklung neuer Hypothesen durch die Möglichkeit, Widersprüche zwischen Ergebnissen entdecken zu können

(5) Verbesserung der Verallgemeinerbarkeit

(6) Präzisere Wirkungsschätzung statt Signifikanzschätzung

(7) Simulation von Replikationen

Zusammenfassungen zu Ergebnissen von Programm-Evaluationen erlauben rationalere bildungspolitische Entscheidungen sowie die Integration tentativer Schlußfolgerungen für die Anwendung theoretischer Forschungsergebnisse für

die Praxis (vgl. LADAS, 1980). Jedoch ist nicht nur die Zusammenfassung im Sinne der Synthese ein gewichtiges Anliegen dieser Integrationstechniken, sondern auch ihr Gegenteil, die Suche nach Erklärungen für die Variation der Ergebnisse, d.h. für die Widersprüche zwischen den Studien (vgl. PILLMER & LIGHT, 1980), LIGHT & SMITH, 1971). Ungeachtet ihrer impliziten und expliziten Ziele bzw. potentiellen Möglichkeiten wurden die verschiedenen Techniken zur Datensynthese klassifiziert. Dem traditionellen Übersichtsartikel (traditional review), den KLAUER (1981) als feuilletonistische Literaturübersicht bezeichnet und inzwischen als qualitative Synthese in die Integrationsliteratur eingegangen ist (vgl. COOK & LEVITON, 1980) stehen die verschiedenen Verfahren auf statistischer Grundlage gegenüber. So definieren PILLMER & LIGHT (1980)

"Data-synthesis will refer here to using formal procedures for combining the results from several empirical studies. We use the word formal to indicate that they are not specific to a particular set of studies. In fact, their systematic nature is their primary strength" (S. 177).

Dabei bilden die "written words of others... 'raw' data for the integrator (FELDMAN 1971, S.87). Entsprechend grenzt GLASS(1976) die Meta-Analyse von der Primär- und Sekundaranalyse ab:

"Primary analysis is the original analysis of data in the research study...
... Secondary analysis is the re-analysis of data for the purpose of answering the original research questions with better statistical techniques, or answering new questions with old data...Meta-Analysis refers to the analysis of analysis. I use it to refer to the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings" (S.3).

Klassifiziert man die unterschiedlichen Verfahren nach der zentral verwandten Statistik, läßt sich folgende Einteilung treffen:

Tabelle 2: Übersicht unterschiedlicher Integrationstechniken

Qualitiative Integration	Quantitative Integration	
traditioneller Übersichtsartikel (Review)	Zählmethode Vorzeichentest Summierte Wahrscheinlichkeit Gemittelte Wahrscheinlichkeit Multiplizierte Wahrscheinlichkeit Addition von Z Addition von t Chi**2 nach JONES & FISKE	Meta-Analyse nach GLASS (Cluster-Approach) BAYES-Ansatz nach FRICKE

1.3.1 Der qualitative Übersichtsartikel

Der Schwerpunkt veröffentlichter Forschungsberichte liegt in der Darstellung individueller Untersuchungsprojekte. Die derzeitig noch häufig angewandte Form, eine Anzahl solcher Studien zu einem Forschungsbereich verbal zusammenzufassen, bietet der qualitative Übersichtsartikel. Obgleich die qualitative Zusammenfassung einen wesentlichen Bestandteil der Kommunikation über Forschungsprojekte und Ergebnisse ausmacht, fehlen Kriterien, die eine systematische Zusammenfassung erlauben. JACKSON (1980) findet bei der Überprüfung von 39 Monographien zur generellen Methodologie in den Sozialwissenschaften vier Hinweise über den Umgang mit veröffentlichten Einzelstudien zum Zweck der Zusammenfassung. Keiner dieser Hinweise umfaßt mehr als zwei Seiten. Der Autor zeigt weiter, daß in 87 Übersichtsartikeln verschiedener Zeitschriften zwischen 1974 und 1975 nur 12 Beiträge Informationen über die angewandten Integrationstechniken liefern. Es verwundert also nicht, wenn die Liste der mit diesem Integrationsverfahren verbundenen Kritiken lang ist. Sie lassen sich zu drei wesentlichen Punkten verdichten.

GLASS (1976) , COOPER & ROSENTHAL (1980)bemängeln die Verzerrung der integrierten Studien durch die subjektive Sichtweise des Zusammenfassers.

"A common method of integrating several studies with inconsistent findings is to carp on the design or analysis deficiencies of all but a few studies - those remaining frequently being one's own work or that of one's student or friends - and then advance the one or two 'acceptable' studies as the truth of the matter" (GLASS 1976, S.4).

FELDMAN (1971) kritisiert die Vernachlässigung der umfangreichen, in den Einzelstudien enthaltene Informationen. Letztlich kritisieren COOPER & ROSENTHAL (1980) die Weichheit der Schlußfolgerungen aufgrund der wenigen in das qualitative Review einfließenden Studien. PILLEMER & LIGHT (1980) schließen daraus, daß die typische Literaturzusammenfassung hoch subjektiv und frei von wissenschaftlichen Standards sei. Die von LADAS (1980) (mit-) geteilte Erfahrung, daß "...the results of such reviews often make it appear that researchers have found little or nothing of value" (S.598) wird inzwischen durch zahlreiche Beispiele belegt (vgl. u.a. GLASS (1976) , KLAUER (1981) , HORNKE (1981) , FRICKE (1978)).

Ein Beispiel
HALL (1978) untersucht mittels Meta-Analyse den Zusammenhang zwischen Geschlechtszugehörigkeit und Genauigkeit in der Dekodierung non-verbaler Kommunikation. Einleitend referiert er die vergangenen Übersichtsartikel zu dieser Fragestellung. Die folgende Tabelle faßt die Ergebnisse zusammen.

Tabelle 3: Übersicht der Review-Befunde nach HALL (1978)

AUTOR(EN)	SCHLUSSFOLGERUNGEN
COLEMAN (1949)	no major sex differences... in the identification of emotions
WEISBERGER (1956)	the results in the literature were not clear
DAVITZ (1964)	the research evidence about sex differences in sensitivity for facial expression is consistently contradictory
TAGIURI (1969)	the literature was somewhat contradictory
EKMAN & FRIESEN (1971)	no differences between male and female subjects were expected and no such differences had been found in the literate of culture data
GITTER, KOZEL, & MOSTOFSKY (1972)	the results of sex analysis in the literature were mixed
MACOBY & JACKLIN (1974)	there is no clear tendency for females to be more sensitive to social cues

Die zitierten Übersichtsartikel spiegeln die Verschiedenheit der Untersuchungsergebnisse ihrer Einzelstudien wider. HALL (1978) integrierte selbst 75 Einzelstudien, die einen mittleren Effektumfang von AM(ES) = .40 zugunsten des weiblichen Geschlechts aufzeigen.

"The analysis revealed that more studies showed female advantage than would occure by chance, the average effect was significantly larger than .00, and more studies reached the p lt .05 level of significance in favor of female than would be expected by chance" (HALL, 1978, S.854).

COOPER & ROSENTHAL (1980) untersuchten, ob in Abhängikeit der Integrationstechnik unterschiedliche Schlußfolgerungen über den jeweiligen Integrationsbereich geschlossen werden. Sie ließen Studenten und Professoren Untersuchungen zum Einfluß des Geschlechtsunterschiedes auf verschiedene Variablen integrieren. Die Ergebnisse zeigen die Analogie zur Demonstration der Integration durch HALL auf. Die statistische Integrationsmethode spiegelt die Ergebnisse der Einzeluntersuchungen präziser wider.

Die Überlegenheit statistischer Integrationsverfahren betonend, meinen COOK & LEVITON (1980), die Kritik an der qualitativen Zusammenfassung betreffe weniger die Methode als vielmehr ihre Handhabung. Bereits vor der Diskussion um die Meta-Analyse wurden in qualitativen Übersichtsartikeln Effektschätzungen verwandt (z.B. bei MISCHEL 1968, CARTWRIGHT 1971, zit. nach COOK & LEVITON). Der Vorwurf, qualititatve Übersichtsartikel seien in ihrer Auswahl siginifikanzorientiert, trifft damit nicht grundsätzlich zu. Ebensowenig schreibt die qualitative Integration vor, nur eine Auswahl von Studien vorzunehmen. Auch sei dieses Verfahren sensitiv genug, Interaktionen aufzudecken. Ähnlich argumentieren LIGHT & Pillemer (1982). Der Primärforschung ähnlich, sei für die Integration von Einzelstudien die Kombination qualitativer und quantitativer Techniken angebracht. Für die Einbeziehung qualitativer Verfahren sprechen sechs Gründe:

(1) Treatments können individualisiert sein; sie gelten nur für bestimmte Personen

(2) kritische Ergebnisse können nur schwer quantitiativ gemesssen werden, z.B. wenn das Ergebnis stark kontext-abhängig ist und die Variablen des Kontextes numerisch schlecht zu erfassen sind

(3) das Treatment weist unterschiedliche Ausprägungen aus, die nicht quantitativ zu erfassen sind

(4) die Unkontrollierbarkeit bei Feldexperimenten, ob das Treatment nur angeboten oder auch aufgegriffen wurde und ob sich die Kontrollgruppe das Treatment anderswertig beschaffen kann

(5) 'falsche' Treatments, daß heißt Treatments, die im Gegensatz zur Mehrheit der Untersuchungen schwache Effekte liefern

(6) der Einfluß qualitativer Deskriptionen auf administrative und politische Entscheidungsträger.

LIGHT & PILLEMER (1982) kann beigepflichtet werden, wenn sie beide Verfahren nicht disjunkt, sondern konjunkt gegenüberstellen. Wie zu zeigen sein wird, besteht eine wesentliche Tätigkeit bei der quantitativen Integration in der Anwendung qualititativer Techniken.

1.3.2 Statistische Integrationsverfahren auf der Basis des kritischen oder empirischen Signifikanzkriteriums

Seit Mitte der 40er Jahre werden verschiedene statistische Integrationsverfahren diskutiert. In Methodenaufsätzen sind diese Verfahren nur isoliert dargestellt, so daß hier eine Übersicht geliefert wird. Eine umfangreiche Explikation verschiedener Ansätze veröffentlicht ROSENTHAL (1978).

Die Zählmethode

Das einfachste Verfahren besteht in der Auszählung (vote counting). Untersuchungen zu einem Gegenstandsbereich werden je nach Ergebnis einer Kategorie zugeordnet. Ein Plus-Zeichen symbolisiert die signifikant positive Beziehung zwischen Prädiktor- und Kriteriumsvariable, ein Minus-Zeichen markiert die signifikant negative Beziehung, eine Null informiert über eine nicht-signifikante Beziehung. Dieses Grundmuster ist unter Umständen erweiterungsfähig, wenn zwischen signifikanten und nicht-signifikanten, aber in dieselbe Richtung weisende Ergebnisse differenziert wird. Die Zelle mit den höchsten Frequenzen liefert dann das Votum für die entsprechende Schlußfolgerung. Eine Vielzahl von Integrationen mit Hilfe dieser Technik liegen vor. LUBORSKY et al. (1971) evaluieren Faktoren, die das Ergebnis von Psychotherapie beeinflussen. Auf der Basis von 166 einbezogenen Studien finden sie, daß "...the largest number deals with patient factors - relatively few with therapist or treatment" (S. 145).

FELDMAN & NEWCOMB (1969) elaborieren diese Methode, in dem sie die Ergebnisse in Abhängigkeit von Probandenmerkmalen zusammenfassen.

Die Schwächen der Auszählmethode wurden wiederholt diskutiert (vgl. GLASS 1976, LIGHT & SMITH (1971), JACKSON 1980), KLAUER (1981), KULIK (1979)). Neben die Kritik, dieses Verfahren überbetone das Signifikanzkriterium, wird besonders herausgestellt, daß dieses Verfahren keine Information über die Größe der Effekte liefert.

"Sind die untersuchten Effekte nur schwach oder mittel, was in pädagogischen Zusammenhängen praktisch immer der Fall ist, so nimmt mit der Zahl der Untersuchungen sogar die Wahrscheinlichkeit ab, daß man den Einfluß nachweisen kann" (HEDGES & OLKIN, 1980, zitiert nach KLAUER 1981, S. 304-305).

Ebensowenig kann der Einfluß weiterer Variablen nicht berücksichtigt werden. Auf methodische Probleme macht JACKSON (1980) aufmerksam. Die Zählmethode unterstellt sowohl einen zweiseitigen Signifikanztest wie dasselbe Signifikanzniveau zwischen den Studien. Es verwundert daher nicht, daß in Abhängigkeit der Integrationsmethode unterschiedliche über-alles-Ergebnisse gefunden werden. DUNKIN & BIDDLE (1974) decken in einer Zusammenfassung keinen Zusammenhang zwischen nicht-direktivem Lehrstil und Schülerleistung auf, während FRICKE (1978) mit Hilfe der Bayes-Statistik einen mittelhohen Zusammenhang auf der Basis 23 integrierter Studien nachweisen kann. HORNKE (1983) stellt ebenso die Ergebnisse verschiedener Integrationen zum Gegenstandsbereich 'vorstruktuierender Lernhilfen' zur Diskussion. Während BARNES & CLAWSON (1975) mit Hilfe der Zählmethode auf der Basis von 32 Einzelstudien die Wirksamkeit vorstruktuierender Lernhilfen verwerfen, zeigen LUITEN et al. (1980) anhand von 135 Studien auf der Basis von Effektgrößen eine mittlere Wirkung von AM(ES) = .21.

Der Vorzeichen- oder Binominaltest

Liegen für N Studien n signifikante Ergebnisse vor, kann die Wahrscheinlichkeit, mit der diese Ergebnisse durch Zufall auftreten, durch die Binominalverteilung berechnet werden (vgl. JONES & FISKE 1953, BROZEK & TIEDE (1952) :

$$P = \sum_{x=n}^{N} \binom{N}{x} P^N q^{N-x}$$

$$= 1 - \sum_{x=0}^{n-1} \binom{N}{x} P^N q^{N-x}$$

Da der Mittelwert einer Binominalverteilung m = N * p und die Standardabweichung SD = SQR(N*p*q) beträgt, schlagen BROZEK & TIEDE (1952) für die Berechnung bei hohem N und einem Mittelwert von m ge 5 die Approximation an die Normalverteilung vor:

$$C.R. = Z = (n - m) / SD$$

Die diesem C.R.-Wert entsprechende Wahrscheinlichkeit informiert darüber, wie hoch die Wahrscheinlichkeit unter der Annahme der Nullhypothese ist, daß n signifikante Ergebnisse von N Studien vorliegen.

Beispiel:
BROZEK & TIEDE (1952) finden in N = 228 Studien 24 Untersuchungen, die mit einer Irrtumswahrscheinlichkeit von p lt .05% signifikant sind.

$$C.R. = \frac{24 - (228 * .05)}{SQR\ (228 * .05 * .95)} = 3.83$$

Die Wahrscheinlichkeit, daß von 228 Studien 24 Studien durch Zufall signifikant sind, beträgt p = .000068.

3.2.3 Summierter Wahrscheinlichkeitswert

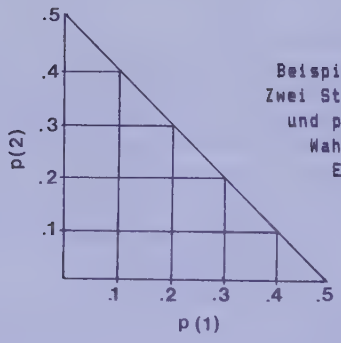
EDGINGTON (1972a) schlägt vor, die Wahrscheinlichkeit, unter der kombinierte Wahrscheinlichkeiten auftreten, nicht multiplikativ, sondern additiv zu bestimmen. Unter Annahme der Nullhypothese sind dabei alle Kombinationen von Wahrscheinlichkeiten als Summe gleichwahrscheinlich. Liefern beispielsweise zwei Studien die Werte p(1) = .2 und p(2) = .3, beträgt die summierte

Wahrscheinlichkeit nach EDGINTON p(s) = .5. Die Wahrscheinlichkeit, mit der eine solche Summe auftritt, beträgt ((Sum p)**2)/2 oder allgemein:

$$P = \frac{(\sum P)^N}{N!}$$

Die Fläche für Summe p = .5 läßt sich geometrisch als rechtwinkliges Dreieck darstellen. Diese Fläche repräsentiert für k = 2 Studien die Wahrscheinlichkeit des gemeinsamen Auftretens und entspricht 1/8 der Fläche, in der alle möglichen Kombinationen von p-Wert-Paaren zwischen 0 und 1 liegen. D.h. die über-alles-Wahrscheinlichkeit einer solchen Summe beträgt p = .125.

Abb. 1: Addition von Wahrscheinlichkeiten



Beispiel:
Zwei Studien mit den Werten p(1) = .2 und und p(2) = .3 liegen vor. Die über-alles-Wahrscheinlichkeit, mit der ein solches Ergebnis auftreten kann, beträgt:
p(k) = ((.2 + .3)**2)/2 = .125
Dies entspricht der Fläche des nebenstehenden Dreiecks.

Da die Summe der kombinierten Wahrscheinlichkeiten größer als 1 werden kann, liefert EDGINGTON (1972a) folgende Korrektur:

$$P_K = \frac{S^K}{K!} - \binom{K}{1}\frac{[S-1]^K}{K!} + \binom{K}{2}\frac{[S-2]^K}{K!} - \binom{K}{3}\frac{[S-3]^K}{K!} + \ldots$$

Die Vorzeichen der einzelnen Terme alternieren. Die Terme werden solange eingesetzt, bis der Wert des Subtrahenten im Zähler größer als S wird. EDGINGTON (1972a) kann aufzeigen, daß die additive Methode gengenüber der multiplikativen Methode bei großer Treatmentwirkung, also bei kleiner Summe der aufaddierten p-Werte, sensitiver ist, während die multiplikative Methode bei großen p-Differenzen vorzuziehen ist.

Die Normalkurvenmethode nach EDGINGTON

Für umfangreiche Einzelstudien empfiehlt EDGINGTON (1972b)die Normalkurvenmethode.

$$Z = \frac{AM(p) - .5}{.2887\sqrt{K}}$$

AM(p) entspricht dem Mittelwert der k-integrierten p-Werte und .2887 ist die Standardabweichung der Population von Wahrscheinlichkeitswerten, unter der Annahme, daß eine kontinuierliche Population eine Standardabweichung von SD = SQR(1/12R) besitzt, wobei R dem Range der Population entspricht. Die theoretische Verteilung der Mittelwerte von Wahrscheinlichkeiten unter der Null-Hypothese hat einen Mittelwert von .5 und eine Standardabweichung von SQR(1/12R)/SQR(k). Für die Werte aus obigem Beispiel ergibt sich:

$$Z = \frac{.218 - .5}{.2887/SQR(5)} = -2,18$$

$$p = .0146$$

Die Wahrscheinlichkeit, daß unter der Annahme der Null-Hypothese ein gemittelter p-Wert von .218 erwartet werden kann, beträg .0146.

Addition von z-Werten

Die einfachste Methode besteht in der Addition der mit den p-Werten assoziierten z-Werte und Division durch die Wurzel der Anzahl kombinierter Studien:

$$Z = \frac{\sum Z}{\sqrt{K}}$$

Jeder z-Wert wird unter der Annahme von Ho als Standardwert betrachtet. Die Varianz der Summe der unabhängigen z-Werte ist die Summe ihrer Varianz, also die Anzahl der Studien, da jede Studie eine Einheitsvarianz aufweist (mü = 0 und sigma = 1).

MOSTELLER & BUSH (1954) erweitern diese Technik, indem sie vorschlagen, jede Standardnormalabweichung zu gewichten. Gewichtungsfaktoren können der der Einzelstudie zugrundeliegende Stichprobenumfang (Freiheitsgrade), aber auch Validitätsgesichtspunkte sein. In diesem Fall wird die Summe gewichteter z-Werte durch die Wurzel der Summe der quadrierten Gewichte dividiert. Im Fall der Gewichtung durch die Freiheitsgrade gilt:

$$Z = \frac{T}{\sigma_T} = \frac{df_1 z_1 + df_2 z_2 + \cdots + df_n z_n}{df_1^2 + df_2^2 + \cdots + df_n^2}$$

Beispiel
Die folgende Tabelle greift die unterstellten Daten zum Geschlechtsunterschied auf:

Tabelle 4: Gewichtete und ungewichtete z-Werte aus fünf Studien

Nr.	Freiheitsgrade	p (einseitig)	z	gewichtetes z
1	40	.12	+1.17	+ 46.8
2	60	.01	+2.33	+139.8 3
3	10	.72	-0.58	- 5.8
4	30	.07	+1.48	+ 44.4
5	20	.17	+ .95	+ 19.0
			+5.35	+244.2

$$Z = \frac{\sum Z}{\sqrt{K}} = \frac{5.35}{\sqrt{5}} = 2.39 \triangleq P = .009$$

$$= \frac{T}{\sigma_T} = \frac{2.44 * 2}{\sqrt{6600}} = 3.03 \triangleq P = .0013$$

Die Gegenüberstellung zeigt, daß der gewichtete Wahrscheinlichkeitswert niedriger ausfällt. Dem niedrigsten p-Wert kommt zugleich aufgrund seiner Stichprobengröße das höchste Gewicht zu.

MOSTELLER & BUSH (1954) modifizieren die Stouffer-Methode, indem sie p-Werte in z-Wert transformieren und darauf einen t-Test für den gemittelten z-Wert berechnen:

$$t = \frac{\Sigma Z / K}{\sqrt{MS_Z / K}}$$

Für die Werte aus der Tabelle 4 ergibt sich:

$$t = \frac{5.35 / 5}{\sqrt{.22513}} = 2.26 \quad , \quad df = 4$$

Analog schlägt ROSENTHAL (1978) die Berechnung von F vor:

$$F = \frac{(\Sigma Z)^2}{N(MS_t)}$$

Die FISHER-Methode

Der Signifikanztest informiert über die Wahrscheinlichkeit, mit der ein beobachteter Datensatz unter der Bedingung der Gültigkeit der Null-Hypothese noch in deren Annahmebereich fällt. FISHER (1948) weist darauf hin, daß bei Vorlage mehrerer Untersuchungen die Einzelwahrscheinlichkeiten keine signifikante Abweichung von der Null-Hypothese anzeigen müssen, daß das "Aggregat jedoch den Eindruck macht, daß die Wahrscheinlichkeiten im ganzen geringer sind, also oft durch Zufall zustande gekommene Wahrscheinlichkeiten" (deutsche Fassung, S. 101). Die Betrachtung der aggregierten Wahrscheinlichkeiten ist über die Chi-Quadrat-Verteilung möglich. Einzelne Chi-Quadrat-Werte sind in Abhängigkeit ihrer Freiheitsgrade wiederum Chi-Quadrat-verteilt:

$$\chi^2 = -2 \lg_e (p)$$

und

$$\chi^2 = \sum -2 \lg_e (P) \qquad df = 2k$$

Das Ergebnis informiert über die Wahrscheinlichkeit, mit dem das Aggregat der Einzelwahrscheinlicheiten durch Zufall auftritt.

Tabelle 5: Datenbeispiel aus FISHER (1948)

df	t	p(einseitig)	-2 log(e) p
6	- .68	.739	.6049
18	1.53	.0717	5.2705
22	2.21	.01932	7.8932
25	1.85	.03809	6.5356
			20.3042

Die Wahrscheinlichkeit, unter der Null-Hypothese einen Ergebnissatz zu erhalten, der einen Chi**2-Wert von 20.3 oder extremer bildet, beträgt weniger als 1% (Chi**2 für df=8 und alpha=.001 = 20.09). Selbst bei zweiseitiger Abweichung ist das Datenaggregat auf dem 2%-Niveau signifikant.

JONES & FISKE (1953) weisen darauf hin, daß die Verwerfung der Null-Hypothese für aggregierte Wahrscheinlichkeitsdaten nicht mit der Zurückweisung der Null-Hypothese für jedes Einzelergebnis identisch ist.

"In combining tests of significance, we are testing a hypothesis about the set as whole, not about the individual members of the set" (S. 380).

Von Vorteil ist, daß in die Integration sowohl die Freiheitsgrade wie die Ergebnis-Richtung eingehen. FISHER (1948 empfiehlt, keine Integrationen für weniger als 15 Untersuchungen nach dieser Methode durchzuführen.

Addition von t-Statistiken

Die Technik nach WINER (1971) besteht in der Summation von t-Statistiken und Division durch den zugehörigen Standardfehler, das sind die Standardverteilungen der jeweiligen t-Verteilungen:

$$\hat{Z} = \frac{\Sigma t}{\sqrt{\Sigma df/(df-2)}} = \frac{\Sigma t}{\sigma_t}$$

Diese Methode basiert auf der Tatsache, daß die Varianz der t-Verteilung df/(df-2) beträgt, also eine Funktion der Freiheitsgrade ist. Das Ergebnis ist dann selbst standardnormalverteilt und an einem bestimmten Wahrscheinlichkeitsniveau bewertbar. Dieses Verfahren kann nur für Studien mit N gt 3 angewandt werden. Liegt der Umfang der Studien unter 10, führt dieses Verfahren jedoch zu keiner guten Approximation an die Standardnormalverteilung.

Die BAYES-Integration nach FRICKE

Während die bisher besprochenen statistischen Integrationsverfahren auf der Zusammenfassung der Wahrscheinlichkeiten von Daten unter der Bedingung der Null-Hypothese basieren, schlägt FRICKE (1977) die Integration von Daten nach dem BAYES-Theorem vor:

$$P[Hi/D] = \frac{p[Hi] * p[D/Hi]}{p[D]}$$

wobei

$$P[D] = \sum P[D/Hi] * P[Hi]$$

Daten könnnen in signifikante und nichtsignifikante klassifiziert werden, so daß für die einzelnen Wahrscheinlichkeiten gilt:

p (signifikante Daten/H(i)) = 1 - beta = eta
p (nichtsignifikante Daten / H(i)) = beta
p (signifikante Daten / Ho) = alpha
p (nichtsignifikante Daten / Ho) = 1 - alpha

Für p(Ho) = p(H(1)) = .5 gilt:

$$P[H_o/s] = \frac{\alpha}{\alpha + \epsilon}$$

Der Cluster-Approach

Der Über-alles-Betrachtung verschiedener Studien auf der Basis von Test-Statistiken stellen LIGHT & SMITH (1971) den Cluster Approach gegenüber. Kombinierte Studien werden nicht als Zufallsstichprobe aus einer gemeinsamen Population, sondern als Cluster betrachtet. Damit ist die Annahme verbunden, daß sich Populationen in jeweils kleinere Teil-Populationen zerlegen lassen, die sich wahrnehmbar und systematisch voneinander unterscheiden. Die Autoren gehen also davon aus, daß die Populationsvarianz nicht als Zufalls- bzw. als Fehlervarianz betrachtet werden sollte, sondern sich aufgrund natürlicher Aggregate in weitere systematische Varianzkomponenten zerlegen läßt:

$$\sigma^2 \text{ (Population)} = \sigma^2 \text{ (C1)} + \sigma^2 \text{ (C2)} + \ldots + \sigma^2 \text{ (Cn)} + e$$

C1 = Cluster 1, Cn = ntes Cluster, e = Fehlervarianz

Cluster lassen sich als "smallest natural unit of the educational process which is available in the data" (S. 446) bilden. Die konkrete Festlegung der Cluster hängt von den jeweiligen Fragestellungen ab.

Beispiel von LIGHT & SMITH (1971)
Die Wirksamkeit von Leseprogrammen für Vorschulkinder soll untersucht werden. Insgesamt liegen in Abhängigkeit der jeweiligen Institution vier verschiedene Programme vor: (i) day care center, (ii) Head Start center, (iii) Montessori und (iv) pre-kindergarten program. Insgesamt untersuchen fünf Studien 12 Klassen.

Tab. 6: Clustermatrix aus LIGHT & SMITH (1971)

Vorschultyp	Studien					
	1	2	3	4	5	Σ
Head Start	A	B	C			3
Day Care	A	B			C	3
Montessori			A	B,C,D		4
Pre-Kindergarten			A		B	2
Σ	2	2	3	3	2	12

Der Tabelle kann entnommen werden, daß Studie (1) jeweils ein Head Start und ein Day Care Programm untersucht (vergleicht), während die Studie (4) drei Montessori-Klassen untersucht. Insgesamt stehen Ergebnisse für 3 Head Start, 3 Day Care, 4 Montessori und 2 Pre-Kindergarten-Programme zur Verfügung. Jede Studie untersucht zumindest zwei Gruppen (Klassen).

Im Gegensatz zu den bisher besprochenen Integrationsverfahren findet die Ergebnis-Zusammenfassung nach dem Cluster-Approach auf der Basis von Rohdaten statt. So existiert für jedes Kind neben dem Post-Test-Wert der Vortestwert und weitere Informationen (Anzahl der Bücher zu Hause, Geschlecht), die als unabhängige Variablen in die Zusammenfassung einfließen können. Absicht der Integration ist es, die Frage zu klären, welches der vier Programme die Lesefertigkeit der Kinder am besten ausformt. Im Gegensatz zur über-alles-Betrachtung sind hier differenzierte Aussagen möglich. So mag beispielsweise ein Vorschul-Typ zu geringen und damit nicht-signifikanten Effekten führen. Die Vortest-Werte zeigen jedoch für Kinder, die diesen Vorschul-Typ besuchen, bereits hohe Werte an, so daß ein Programmeffekt erwartungsgemäß niedrig ausfällt. Die über-alles-Interpretation auf der Basis von Teststatistiken würde zu der statistisch zwar richtigen, aber inhaltlich unkorrekten Aussage gelangen, daß dieser Programm-Typ 'im Ver-

gleich zu anderen Programmen' die Lesefertigkeit nicht erhöht, ohne den Selektionseffekt zu berücksichtigen.

LIGHT & SMITH nennen insgesamt fünf Merkmale, durch die sich Cluster unterscheiden können:

(1) Mittelwert-Unterschiede
Natürliche Aggregate unterscheiden sich u.U. durch Selektionsprozesse bereits in ihren Vortest-Werten, so daß eine über-alles-Betrachtung zu falschen Schlußfolgerungen führen können (siehe oben). Erst wenn sich die Vortest-Nachtest-Vergleiche innerhalb einer Vorschulinstitution nicht (signifikant) unterscheiden, ist es sinnvoll, die Cluster zusammenzufassen.

(2) Varianzunterschiede
Die Clustervarianzen können sich unterscheiden. Angenommen, zwei Programme führen zu denselben Nachtestwerten, jedoch mit unterschiedlichen Varianzen. Vor der Zusammenfassung sollte versucht werden, diese Unterschiede zu erklären.

(3) Kovarianz-Betrachtung
Die Cluster können sich in ihrer Beziehung zwischen der abhängigen Variablen und eventuellen Kovariablen unterscheiden. So mag der Programmeffekt innerhalb eines Institutions-Typs unterschiedlich ausfallen, weil bestimmte Kovariablen unterschiedlich ausgeprägt sind.

(4) Personen-mal-Treatment-Interaktion
Personenmerkmale, die mit einem Treatment kovariieren, können ebenfalls die Zusammenfassung verfälschen. Führt beispielsweise die Geschlechts-Variable zu unterschiedlichen Ergebnissen innerhalb einer Institution und sind Cluster unterschiedlich von Jungen und Mädchen frequentiert, können sich im Extremfall Effekte aufheben oder zumindest stark verringern, solange diese Interaktion unberücksichtigt bleibt.

(5) Kontext-Effekte
Kontext-Variablen liefern weitere Quellen möglicher Ergebnisvariationen sowohl zwischen wie innerhalb der Programme. Hierzu können Merkmale des Programm-Anwenders (hier der Lehrer) wie die Klassengröße gehören. Solche Einflüsse müssen entdeckt und herauspartialisiert werden können.

LIGHT & SMITH (1971) begründen den Cluster-Approach mit der Gefahr, durch über-alles-Aussagen könnten wichtige Unterschiede verloren gehen bzw. verwischt werden. Sie sensibilisieren zu Recht auf verschiedene Einflußquellen, vor deren Hintergrund Ergebnisse reflektiert werden müssen. Wie zu zeigen sein wird, bedarf es hierzu nicht notwendigerweise Rohdaten.

Kritik an den Über-alles-Wahrscheinlichkeitsschätzungen

Die Integration verschiedener Studien erfordert eine gemeinsame Forschungshypothese sowie weitgehende methodologische Kongruenz der Studien. Mit der über-alles-Schätzung wird dies suggeriert. Das Gegenteil bestimmt jedoch den wissenschaftlichen Alltag: Studien unterscheiden sich nicht nur in der operationalen Realisation diverser theoretischer Konstrukte, sie unterscheiden sich vielmehr auch in den Merkmalen ihrer Probanden, in der methodischen Anlage der Studien, in den Auswertungs- und Adjustierungsverfahren, um nur wenige Punkte zu nennen. ELASHOFF & SNOW (1971) bemängeln dann auch, daß diese Form der Zusammenfassung grundlegende Unterschiede zwischen den Studien verdeckt. Diese Unterschiede, so HILGARD (1978) können nicht als Zufallsfehler unberücksichtigt bleiben.

"... it can't be assumed that other sources of variation in a large number of related studies will be random and can be treated as a background of noise against which the influence of the selected variables stands out clearly" (S. 395).

Zudem gehen alle unter dem Aspekt der Signifikanz und Replikation genannten Probleme in die über-alles-Schätzung mit ein. ELASHOFF & SNOW (1971) verweisen auf die Abhängigkeit des empirischen p-Wertes von den Anforderungen, die an das einzelne Testverfahren gestellt sind. Sie vermuten, daß der wahre p-Wert häufig viel höher liegt als der beobachtete.

LIGHT & SMITH (1971) sehen in der Integration gerade die Aufgabe, auf jenseits der Forschungshypothesen liegende Varianzquellen aufmerksam zu machen. Narrative Überblicke, wie sie das qualitative Review oder die Signifikanz-Integration darstellen, liefern keinen umfassenden und befriedigenden Beitrag.

1.4. Die Meta-Analyse

Unter Meta-Analyse werden seit GLASS (1976) statistische Integrationsverfahren subsummiert, die über die Wahrscheinlichkeitsschätzung hinaus den Einfluß unterschiedlicher Variablen auf das Ergebnis zu bestimmen versuchen.

1.4.1 Grundrationale der Meta-Analyse

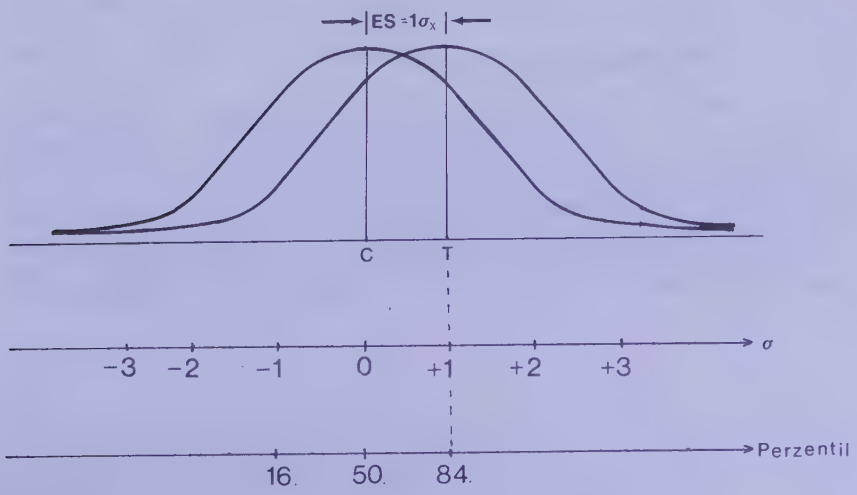
Die meta-analytische Integration erlaubt die Synthese von Einzelstudien unter derselben Fragestellung bzw. zu demselben Phänomenbereich auf der Basis der in den Studien enthaltenen Statistiken. Die zentrale Statistik der Meta-Analyse bildet die Effektgröße (effect size): für eine gegebene Studie wird aus den vorliegenden - veröffentlichten - Statistiken die mittlere Differenz der abhängigen Variablen zwischen Untersuchungs- und Kontrollgruppe, dividiert durch die Fehlervarianz (Standardabweichung), berechnet:

$$ES\ (\text{Effektgröße}) = (AM(T) - AM(C)) \,/\, SD(C)$$

AM(T) = Mittelwert der Treatmentgruppe, AM(C) Mittelwert der Kontrollgruppe, SD = Standardabweichung (der Kontrollgruppe).

Die Standardisierung durch die Fehlervarianz erlaubt den Vergleich unterschiedlicher 'absoluter' Mittelwerte. Es wird auf die Standardnormalverteilung (mü = 0 und sigma = 1) transformiert. Für eine Studie kann bei Vorlage der Mittelwerte und Standardabweichung(en) beispielsweise eine Effektgröße von ES = + .5 zugunsten der Treatmentgruppe berechnet werden:

Abb. 2: Graphische Darstellung der Effektgröße auf der Grundlage der Verteilung der Gruppenwerte



Bei einem Ergebnis von ES = .5 sigma kann dann gesagt werden, daß im Mittel von einem dem Treatment ausgesetzten Probanden erwartet wird, 'besser' abzuschneiden als 70% der nicht mit dem Treatment konfrontierten Mitglieder der Kontrollgruppe. D.h., es wird für die jeweilige Effektgröße das entsprechende Perzentil abgelesen.

Die aus den einzelnen Studien errechneten Effektgrößen können nun in Abhängigkeit verschiedener Fragestellungen, Variablen und/oder Meßinstrumente geblockt, gemittelt und verglichen werden. Je nach Qualität der Daten und des Umfangs der aus den Einzelstudien kodierbaren Informationen können spezifische Einflüsse auf die Effektgrößen beschrieben werden. Mit Hilfe der multivariaten Regressionsanalyse ist es möglich, den gemittelten Effekt um Einflüsse aufgrund unterschiedlicher Stichproben-, Designmerkmale, Personenmerkmale oder Merkmale der Experimentatoren etc. zu adjustieren. Im einzelnen besteht die Durchführung einer Meta-Analyse aus folgenden Teilschritten:

1. Generelle Forschungsfrage
2. Literatur-Recherche
3. Spezifikation der Forschungsfragen
   und Erst-Selektion der Einzelstudien
4. Kodierung der Einzelstudien
5. Berechnung oder Schätzung der Effektgrößen
6. Berechnung der mittleren Effekte in Abhängigkeit
   der Fragestellungen und Varianzaufklärung und
   Explorative Datenanalyse
7. Darstellung und Interpretation der Ergebnisse

Beispiel: Psychotherapie-Studie von GLASS

GLASS (1976) berechnet auf der Grundlage von 375 Einzelstudien zur Wirkung von Psychotherapie eine mittlere Effektgröße von AM(ES) = .68. Demnach schneidet im Mittel ein Proband, der an einer Psycho-Therapie teilgenommen hat, besser ab als 75.5% der Mitglieder der Kontrollgruppe, d.h. derjenigen, die nicht an einer Therapie teilnahmen. Diese Studien blockte GLASS in Abhängigkeit der angewandten Therapie-Verfahren. Zwei Superklassen konnten mit Hilfe multidimensionaler Skalierungsverfahren gebildet werden, die 'behaviorale' und die 'non-behaviorale' Therapieform. Für verhaltensthera-

peutisch behandelte Probanden zeigt sich eine mittlere Effkgröße von AM(ES) = .69 sigma und für nicht-verhaltenstherapeutisch behandelte Probanden ein Effekt von AM(ES) = .62 sigma. Die Unterschiede zwischen beiden Therpieverfahren sind marginal. Blockt man die Ergebnisse in Abhängigkeit der abhängigen Variablen (Krankheitssymptom), zeigt dieselbe Studie (GLASS & SMITH, 1977) für die Behandlung von Furcht bzw. Angst eine mittlere Effektgröße von AM(ES) = .97 sigma, während der Wert für die Behandlung von Schul- und Arbeitsleistung bei AM(ES) = .31 sigma lag. Die Autoren zeigten weiter, daß 25% der Varianz in den Ergebnissen durch die Spezifikation weiterer Einflußquellen (u.a. Intelligenz der Probanden, Erfahrung des Therapeuten) reduziert werden konnte.

1.4.2 Literaturrecherche

Idealiter gehen in die Meta-Analyse alle Studien zu einem Gegenstandsbereich ein. Dabei stellen sich eine Reihe pragmatischer Probleme:

(i) nicht alle Studien sind während eines kalkulierbaren Zeitraums verfügbar
(ii) nicht alle Studien liefern hinreichend Informationen, die für die Meta-Analyse benötigt werden
(iii) die unabhängigen und/oder abhängigen Variablen stimmen nur wenig mit den in Frage stehenden Konstrukten überein
(iv) Die Analyse der Einzelstudien ist zeitaufwendig; genügend Arbeitsressourcen müssen zur Verfügung stehen.

zu (i): Verfügbarkeit über Einzelstudien
Dokumentationssysteme sind für die Literaturrecherche unerläßlich. GLASS (1978b) weist in diesem Zusammenhang auf ein fehlendes System regelmäßiger Information hin:

"Educational research lacks a regular compendium of abstract of the archival literature with backward and forward citations" (S. 353).

KULIK, KULIK, & COHEN 1980 nennen die folgenden systematischen Suchquellen:

"(a) Compendex, the machine-readable version of the Engineering Index
(b) Comprehensive Dissertation Abstracts
(c) ERIC, a data base on educational materials from the Educational Research Center, consisting of the two files Research in Education and Current Index to Journals in Education
(d) Inspec, an on-line file correspondending to the printed Physics Abstracts, Elecric and Electronic Abstracts, and Computer and Control Abstracts

(e) NTIS, the data base of the National Technical Information Service, consisting of reports on government sponsored research, development, and engineering
(f) Psychological Abstracts
(g) Scisearch, a multidisciplinary index comprised of all the records published in Science Citation Index; and
(h) Social Science Citation Index...
The bibliographies in articles located through computer searches provided a second source of studies for meta-analysis" (S. 528).

zu (ii) Unzureichende Information
Nur ein geringer Prozentsatz der auffindbaren Literatur fließt tatsächlich in eine Meta-Analyse ein. Je nach Umfang und Strenge der angesetzten 'Filter' scheiden solche Studien aus, die den Aufnahmekriterien nicht entsprechen: die Studien liefern entweder unzureichende Statistiken, so daß die Schätzung des Effekts nicht möglich ist, die unabhängigen oder abhängigen Variablen weichen zu weit von dem angelegten Konstrukt ab, oder die methodologischen Mängel sind zu groß als daß die Untersuchungen in die Meta-Analyse einfließen sollten. Die nachfolgende Tabelle zeigt an einer zufälligen Auswahl von durchgeführten Meta-Analysen den Umfang der erstgefunden Studien, die Auswahlkriterien und die Anzahl tatsächlich aufgenommener Studien.

Berücksichtigt man die Extreme nicht, so wurden pro Meta-Analyse im Durchschnitt 720 Studien gefunden. Davon verblieben nach einer ersten Selektion im Mittel 242 Studien, die die Grundlage für den weiteren Integrationsprozeß lieferten. Tatsächlichen gingen in die jeweiligen Meta-Analysen im Mittel nochmals nur die Hälfte (im Mittel 121 Studien auf der Basis der vorliegenden Tabelle) ein.

zu (iii): Arbeitsintensität
Die Durchführung einer Meta-Analyse ist ähnlich einer Originaluntersuchung äußerst arbeitsintensiv. Dies deuten bereits die oben aufgezeigten Zahlen über den Input in die Anayse an. Weiter unten wird die Analyse der Einzelstudien vorgestellt. Es bedarf hierzu der Entwicklung eines entsprechenden

Tab. 7: Ausfallquoten zwischen erster und zweiter Literaturselektion und in eine Meta-Analyse einfließender Einzelstudien

Autor(en)		Gegenstandsbereich	erste Auswahl (= 100%)		Selektionskriterien	nach der Selektion	Studien mit hinreichenden Daten
COHEN (P.A.)	1981	Schülerselbsteinschätzung und Schülerleistung	450	7	(1) Daten aus Feldstudien (2) Klassen als Analyseaggregat (3) spezifische UV- und AV-Anforderungen	105 23%	41 9%
COHEN, EBELING, & KULIK (J.A.)	1981	visuelle Instruktion konventionellem Unterricht (UV); Schülerleistung (AV)	1400	11	(1) Daten aus Feldstudien (2) spezielle UV- und AV-Anforderungen (3) frei von groben methodischen Mängeln	320 23%	72 5%
KULIK, KULIK & COHEN	1980	Computerunterstützter gegenüber tradotionellem Unterricht (UV); Schülerleistung (AV)	500	17	wie COHEN, EBELING & KULIK	180 36%	59 12%
KULIK, COHEN, & EBELING	1981	programmierter versus traditionellem Unterricht (UV); Schülerleistung (AV)	5000	21	wie COHEN, EBELING & KULIK	624 12.5%	56 1.1%
LUITEN, AMES, & ACKERSON	1980	Struktuierte Lernhilfen (UV); Lernen und Behalten (AV)	170	25	keine Angaben	-	135 79%
PFLAUM, WALBERG, KARESIANES, & RASHER	1980	Lese-Unterricht	665	28	Zufallsauswahl	199 30%	97 14.6%
SMITH & GLASS	1977	Psychotherapie-Verfahren	1000	32	(1) Kontroll- und Treatmentgruppe (2) Definition von Psychotherapie nach MELZOFF & KORNREICH (1970)	500 50%	375 37.5%
GLASS & SMITH	1979	Klassengröße (UV); Schülerleistung (AV)	300	37	keine Angaben	150 50%	80 26.7%
KLAUER	1981	Zielorientiertes Lehren und Lernen	226	40	Untersuchungsanordnung	-	16 7%
DUBIN & TAVEGIA	19	verschiedene Lehrmethoden	95	43	Untersuchungsanordnung	74 78%	18 19%

Kodiersystems. Dieses Kodiersystem sollte einer Reliabilitätsüberprüfung ausgesetzt werden. Die interpersonale Reliabilitätsüberprüfung liefert Hinweise auf die Genauigkeit der die Information abfragenden Items. Die intrapersonale Reliablitätsüberprüfung kontrolliert Lernprozesse während des Kodiervorgangs. Genauere Hinweise zum Umgang mit diesem Schritt der Meta-Analse liegen nicht vor. GLASS (1978b) schätzt allein für die Kodierung pro Effektgröße ca zwei Stunden, was bei durchschnittlich 300 Effektgrößen pro Meta-Analyse allein mit 600 Arbeitsstuden zu Buche schlägt. Das Erst- und Zweitlesen allein der Untersuchungen, die nicht in die Analyse einfließen, wurde dabei ebensowenig berücksichtigt, wie die Suche und Beschaffung der Untersuchungen.

1.4.3 Entwicklung von Forschungsfragen und Erstselektion der Individualstudien

Herkömmliche Integrationsverfahren begnügen sich mit der Summation von Ergebnissen: sie beschreiben den Forschungskorpus zu einem bestimmten Gegenstandsbereich, d.h. die Integratoren stellen "Was ist der Fall-Fragen" (vgl. STEGMÜLLER 1964). JACKSON (1980) zeigt auf, daß für herkömmliche Synthesen die Auswahl von Fragen an den Forschungsbereich oder die Konfrontation mit Hypothesen ohne Wichtigkeit war. FELDMAN (1971) betont, daß Primärforschung und Integration sich hinsichtlich ihrer Zielsetzung nicht unterscheiden sollten. Aufgabe der Integration ist es, Hypothesen zu testen, neue Informationen zu generieren, größere Verallgemeinerungen zu erlauben und neue Theorien zu entwickeln. Um dem 'Typ-III-Fehler' zu begegnen - nämlich richtige Antworten auf falsche Fragen zu erhalten (vgl. LIGHT & PILLEMER 1982) - nennt JACKSON (1980) als erste Aufgabe einer Meta-Analyse die Entwicklung von Fragen bzw. Hypothesen.

"Skill in asking good questions is probably as important to the progress of science as skill in finding tentative answers or verifying those answers" (S. 442).

Er unterscheidet zwischen 'Kausalfragen' (Wirkt X auf Y?) und Fragen nach der Variation (Wirkt X auf Y nur unter der Bedingung Z?). Die letzte Frage ist dann von besonderer Wichtigkeit, wenn Ergebnisse mehrerer/vieler Untersuchungen Inkonsistenzen aufweisen. Vier Quellen sollten nach JACKSON (1980) herangezogen werden, um Fragen entwickeln zu können, d.h., den Typ-III-Fehler gering zu halten:

(1) verfügbare Theorien
(2) vorliegende Übersichtsartikel
(3) die zur Verfügung stehenden Einzelstudien
(4) Intuition und Einsicht (S. 443).

Während die Mehrheit der zwischen 1977 und 1980 durchgeführten Meta-Analysen unterschiedliche Wirkungen verschiedener programmatischer Variablen

beschreiben, unterlegen einige Datensynthesen erklarungheischende Warum-Fragen' (STEGMÜLLER, 1974). WALBERG und Mitarbeiter (HAERTEL & WALBERG 1980, UGUROGLU & WALBERG 1980, IVERSON & WALBERG 1980, WILLIAMS, HAERTEL & WALBERG 1980, PARKERSON, HAERTEL & WALBERG 1980, HARTEL, WALBERG & HAERTEL 1980, LYSAKOWSKI & WALBERG 1980, SCHILLER, WALBERG & HAERTEL 1980, PLAUM et al. 1980, WAXMAN & WALBERG 1980, KREAMER et al. 1980, HAERTEL, HEARTEL & WALBERG 1982, HEARTEL et al. 1979, IVERSON & WALBERG 1979) liefern eine systematische Analyse von neun Konstrukten (Alter/Entwicklungsstand, Fähigkeit, Motivation, Qualität und Quantität der Instruktion, häusliches Milieu, Peer-Einfluß, Sozialklima der Klasse, Massenmedien), deren Einfluß auf die Leistung von Schülern sie annehmen. Ihr Ziel ist es, die theoretischen Rationale herauszufinden, die einem solchen Modell unterliegen. HUNTER und Mitarbeiter (SCHMIDT & HUNTER 1977, HUNTER & SCHMIDT 1979, SCHMIDT et al. 1979) entwickeln und überprüfen ein Modell zur Validierung von 'Employment-Tests'. Mit Hilfe der meta-analytischen Integration kommen sie zu dem Schluß,

"that the most of the between study variation in observed validity coefficients of various aptitude and ability-tests for two typs of clerical work and for the job of first-line supervisors is artifactual in nature. These results cast additional doubt on the traditional belief that employment test validities are situationally specific" (SCHMIDT et al. 1979, S.257.

SCHWAB et al. (1979) synthetisieren die Ergebnisse der Untersuchungen zur VROOM'schen Theorie zwischenmenschlicher Erwartungen (VROOM, 1964). Sie entwickeln solche Fragen, die die Bewährung von Theorien prüfen. LADAS (1980) demonstriert die Bedeutung erklärungssuchender Fragen, wenn Widersprüche zwischen den Ergebnissen auftreten. Mit Hilfe der Vote-Counting-Methode erfaßt er Untersuchungen zur Frage "Does the process of note-taking itself aid recall?". 17 Studien liefern ein sigifikantes Ergebnis, 16 Studien weisen keine Signifikanz aus. Für diese Studien entwickelt der Autor Fragen, die alternative Erklärungen für diese Ergebnisse abprüfen sollen.

Das erste Ensemble von Fragen wird mit dem recherchierten Literaturkörper konfrontiert. Dabei fallen solche Studien heraus, die aufgrund ihres spezifischen Untersuchungsbereiches nicht in den Antwortbereich fallen.

1.4.4 Analyse der Einzelstudien

Für die Meta-Analyse wird angenommen, daß sich der gemittelte Effekt abhängig von 'relevanten' Merkmalen bilden läßt, oder daß die Varianz der gemittelten Effektgrößen in Abhängigkeit weiterer Einflußfaktoren aufgeklärt werden können. Dies bedeutet die sorgfältige Entwicklung eines Analyserasters, das auf jede Untersuchung gelegt wird, um die für die Fragestellungen wichtige Information herausfiltern zu können. Bei einer Vielzahl durchgeführter Untersuchungen wird neben den gegenstandsspezifischen Variablen Information für folgende Kategorien extrahiert:

(a) Methodologie, insbesondere Gültigkeitsapekte, Präzision der publizierten Angaben, Stringenz der Fragestellung sind die wesentlichen Aspekte der zu kodierenden Information. GLASS (1976) arbeitet mit einer Ratingskala zur Determination der über-alles-Qualität jeder Einzelstudie, während COHEN (1981) zur Kontrolle der internen und externen Gültigkeit zwölf Items entwickelte.

(b) Treatmentspezifikation: neben dem Treatment werden Bedingungen seiner Applikation (wie Dauer, Merkmale des Experimentators) und weitere Begleitumstände kodiert

(c) Populationsmerkmale bzw. Stichprobenmerkmale: soweit möglich und sinnvoll, werden die Merkmale der untersuchten Personen aufgenommen; dies können neben dem Geschlecht, sozioökonomischer Status, Schulbildung etc, die Ergebnisse unterschiedlicher begleitend erhobener Testergebnisse sein

(d) Meßinstrument: Nur selten stimmen über Untersuchungen hinweg die Meßinstrumente überein. Zwar erlaubt die Standardisierung der Mittelwertdifferenz eine skalenfreie Betrachtung der Ergebnisse, dennoch kann der systematische Einfluß der Effektgrößen in Abhängigkeit vom Meßinstrument zwischen gemittelten Effektgrößen nicht ausgeschlossen werden. Dies wird in der Untersuchung von WHITE (1976) deutlich. In 143 Studien wurden zur Bestimmung des sozioökonomischen Niveaus über 70 verschiedene Variablen verwandt. Er bildet daher ein vierkategoriales System, mit dem er die Effektgrößen mittelte.

(e) Publikationsmerkmale: Durch die erste Meta-Analyse von GLASS (1976) hat sich die Tradition herausgebildet, Publikationsmerkmale mitzukodieren. Dahinter verbirgt sich einerseits die bereits dargestellte Diskussion um die Verzerrung von Forschungsergebnissen in Abhänigkeit der Publikationsart, zum anderen kann natürlich mit Hilfe dieser Daten der Forschungsprozeß nachgezeichnet werden.

Zur Analyse und Bewertung von Studien liegen verschiedene Prüfsysteme vor. Sie erlauben die kritische Reflexion insbesondere der methodischen Anlage dieser Untersuchungen und die Überprüfung alternativer Erklärungsmöglichkeiten (JOHNSON Jr.,1957;FARQUHAR & KRUMBOTZ, 1959; STEPHENS, 1967, WOLF, 1973; YI MING HSU, 1979, STRAUSS, 1969). Das Ziel einer kritischen Analyse besteht nach STEPHENS (1967) darin, daß

"...the critical reader of a published investigation should be in a better position to decide whether or not the investigation provides dependable evidence that 'a change in x plays some part in producing change in y' " (S. 102).

JOHNSEN Jr. (1957) schlägt sechs Kategorien vor, die die Evaluation von Forschungsergebnissen mit Hilfe einer Rating-Skala erlaubt. In einer Erprobungsstudie konnte er zeigen, daß die Übereinstimmung der Rating-Urteile zwischen Studenten bei r = .75 und zwischen Experten bei r = .79 lag. FARQUHAR & KRUMBOLTZ (1959) liefern fünf Kategorien mit jeweils 2 bis 7 Unterfragen. Hinweise zur Beurteilung der extrahierten Information liefern sie nicht. STEPEHENS (1967) entwickelt demgegenüber ein Flußdiagramm, mit dessen Hilfe die in den Untersuchungen genannte Information durch rivalisierende Hypothesen kontrastiert werden kann. Umfangreiche Aspekte, die bei der Planung einer Untersuchung zu berücksichtigen sind, listen BORG & GALL (1973) anhand von 123 Fragen auf, die 12 Kategorien zugeordnet sind. Tabelle 8 zeigt die verschiedenen Prüfverfahren in einem Überlick.

Tab. 8: Darstellung verschiedener Check-Systeme in der Übersicht

JOHNSONS (1937)	FARQUHAR & KRUMBOLTZ (1959)	STEPHENS (1967)	STRAUSS (1969)	BORG & GALL (1973)	WOLF (1973)	GLASS (1) (1977)	YHI MING HSU (1979)
PROBLEMSTELLUNG	PROBLEMSTELLUNG	liefert ein auf JA - NEIN - Alternativen hinauslaufendes Entscheidungsschemas	PROBLEMSTELLUNG	UMFANG UND DEFINITIONSBEREICH DER STUDIE	RELEVANZ DER UNTERSUCHUNG	TREATMENTAPPLIKATION	FORSCHUNGSPROBLEM
MATERIALIEN	DESIGN	Mittels expliziter und alternativer Hypothesen wird gefragt, ob X eine Veränderung in Y bewirke oder ob andere Faktoren (mit-) wirken.	FORSCHUNGSTRADITION	HYPOTHESEN	FRAGESTELLUNG und HYPOTHESEN	KLIENTEN- und THERAPEUTENMERKMALE	REFERENZLITERATUR
PROBANDEN	METHODISCHE VORGEHENSWEISE		HYPOTHESEN	HINTERGRUNDANNAHMEN	FORSCHUNGSSTRATEGIE	MESSINSTRUMENT	FORSCHUNGSDESIGN
METHODISCHE VORGEHENSWEISE	ANALYSE		HINTERGRUNDANNAHMEN	DEFINITIONEN	MESSINSTRUMENTE	INTERNALE VALIDITÄT	DATENAUSWAHL und -ANALYSE
ERGEBNISSE	INTERPRETATION		POPULATION	METHODE	GRUNDGESAMTHEIT - STICHPROBE	FORM DER PUBLIKATION	SCHLUSSFOLGERUNGEN und VERALLGEMEINERUNGEN
SCHLUSSFOLGERUNGEN			STICHPROBE	DESIGN	DURCHFÜHRUNG DER MESSUNG - DATENSAMMLUNG		STIL und ORGANISATION DES BERICHTS
			FORSCHUNGSINSTRUMENTE	STICHPROBENVERFAHREN	DATENAUFBEREITUNG		
			UNTERSUCHUNGSPLAN	PERSÖNLICHKEITSMERKMALE	STATISTISCHE AUSWERTUNG		
			DURCHFÜHRUNG	MESSMETHODEN	ENTSCHEIDUNG UND INTERPRETATION DER ERGEBNISSE		
			VORSICHTSMASSNAHMEN	STATISTISCHER UMGANG MIT DEN ERGEBNISSEN	VERALLGEMEINERUNG		
			ROHDATEN	BEURTEILUNG DER STUDIE			
			ERGEBNISSE				
			STATISTIK				
			INTERPRETATION				
			SCHLUSSFOLGERUNG				
			GRENZEN DER UNTERSUCHUNG				
			NEUE PROJEKTE				
			VERBESSERUNGEN				

(1) streng genommen hat GLASS kein Kategoriensystem zur Analyse von Untersuchungen entwickelt; sein Check-System bezieht sich auf die von ihm durchgeführte Meta-Analyse

1.4.5 Berechnung und Schätzung der Effektgrößen

Die Berechnung der Effektgröße impliziert eine Fülle von Fragen. Die Differenzbildung allein auf der Grundlage der Nachtestwerte (AM(T)/NT) - AM(C)/NT) läßt die Differenzierung im Vortest unberücksichtigt bzw. suggeriert AM(T)/VT = AM(C)/VT = AM(C)/NT. HEDGES (1978) fragt darüber hinaus, ob mit der Effektgröße die Beschreibung eines standardisierten Effektumfangs oder der Schätzer für Populationsparameter vorliegen. Dies würde unterschiedliche Basisannahmen implizieren. Ebensowenig ist bisher eindeutig geklärt, welche Modellannahme bei der durch das Treatment vermutlich hervorgerufenen Veränderung der Standardabweichung der Treatmentgruppe herangezogen werden muß. Weiter unterstellt die Berechnungsformel für die Effektgröße ein idealit echt-experimentelles Design (vgl. CAMPBELL & STANLEY (1963)). COOK & CAMPBELL (1979) zeigen jedoch die verschiedensten Vortest-Nachtestzusammenhänge auf, die im Rahmen experimenteller Anordnungen auftreten können.

Die Berechnung der Effektgöße in Abhängigkeit der Votestdaten

McBAW & GLASS (1980) verweisen auf unterschiedliche Möglichkeiten zur Berechnung der Effektgröße. Unterscheiden sich die Mittelwerte der Vortestdaten zwischen Treatment- und Kontrollgruppe, empfehlen sie die Berechnung eines Veränderungswertes:

$$G(T) - G(C) = (AM(T)/NT - AM(C)/NT) - (AM(T)/VT - AM(C)/VT$$

Eine weitere Möglichkeit bietet bei Vortestdifferenzen die Berechnung der Residuen:

$$g(T) - g(C) = (AM(T)/NT - AM(C)/NT) - by.x\ (AM(T)/VT - AM(C)/VT)$$

Die durch Kovarianz adjustierten Werte erlauben die verzerrungsfreie Betrachtung der Nachtestdifferenzen.

Der Standardisierungskoeffizient

Als Standardisierungskoeffizient schlagen GLASS und Mitarbeiter (GLASS (1976), GLASS & SMITH (1977), McBAW, GLASS & SMITH (1981)) die Standardabweichung der Kontrollgruppe vor. Erstens kann durch diese Transformation die relative Treatmentwirkung in Bezug auf die Nichtapplikation oder in Bezug auf ein alternatives Treatment in Begriffen der Verteilung von Werten nicht- bzw. anders behandelter Personen beschrieben werden (... ist im Mittel besser als...). Zweitens erlaubt die Anwendung der Standardabweichung der Kontrollgruppe die Berechnung vergleichbarer Effektgrößen dort, wo in Einzelstudien verschiedene Treatments im Vergleich zu derselben Kontrollgruppe untersucht werden.

Beispiel nach GLASS
In einer Studie werden zwei unterschiedliche Treatmens untersucht. Die folgenden Ergebnisse liegen vor:

	Treatment I	Treatment II	Kontroll-gruppe
AM	50	50	48
SD	10	1	4

Bei Verwendung der Standardabweichung der Kontrollgruppe beträgt die Effektgröße für beide Treatmentgruppen im Vergleich zur Kontrollgruppe ((50 - 48) / 4 = .5 sigma). Wird die Standardabweichung der jeweiligen Treatmentgruppe verwandt, liegt die Effektgröße für die Treatmengruppe I bei ((50 - 48) / 10 = .2 sigma) und für die Treatmentgruppe II bei ((50 - 48) / 1 = 2.0 sigma). Es werden Unterschiede erzeugt, die nicht auf die Mittelwertdifferenzen zurückzuführen sind. Dies ist nach GLASS (1978) Grund genug, die Standardabweichung der Kontrollgruppe anzuwenden (vgl. S. 371).

HEDGES (1979) wirft die Frage auf, ob die Effektgröße als 'Rohwert' für beschreibende oder für Inferenzstatistiken verwandt wird. Bilden die in eine Meta-Analyse einfließenden Studien die Stichprobe einer hypothetischen Population, so gilt:

$$\text{delta} = (\ddot{\mu}(T) - \ddot{\mu}(C) \ / \ \text{sigma})$$

Währen die Gruppenmittelwerte die besten Schätzer für die Populationsmittelwerte bilden, gibt für die Varianz der Population die Varianz innerhalb der Gruppen den besten Schätzer ab:

$$\hat{\sigma} = SD_p = \sqrt{\frac{(n_1-1)SD_1^2 + (n_2-1)SD_2^2}{n_1 + n_2 - 2}}$$

Für jede Einzelstudie gilt demnach:

$$\text{delta(geschätzt)} = (AM(T) - AM(C) \ / \ SD(p)$$

Damit begründet HEDGES die Verwendung einer Effektgröße in Analogie zu COHEN's (1966) 'd'-Wert. Desweiteren führt HEDGES (1980) die Adjusierung des geschätzen Delta-Wertes ein. Für Delta liegt dann eine gute Schätzung vor, wenn er sowenig wie möglich verzerrt ist und möglichst wenig Varianz aufweist.

"The bias for an estimator is therefore the degree to which the estimator consistently over or under estimates the parameter of interest. Unbiased estimates do not err consistently in either direction. The variance of an estimator is a measure of the degree of variability of estimates across samples. Thus, if two estimators have comparable bias, the one with the smaller variance would usually be prefered because the estimates obtained from it are more consistent across samples" HEDGES (1980), S. 26)).

Wenn angenommen werden kann, daß die Werte innerhalb eines jeden Experiments mit gleicher aber unbekannter Varianz normal verteilt sind, können Bias und Varianz bestimmt werden. Unter diesen Design-Bedingungen folgt 'g' einer nicht-normalen, nicht-symmetrischen und schiefen Verteilung für den Fall, daß Sigma von Null verschieden ist. HEDGES bietet die Korrektur

$$g(u) = Cm * g$$

an, wobei Cm der Korrekturfaktor ist, wenn die Standardabweichung m Freiheitsgrade aufweist:

$$Cm = \Gamma(m/2) / (\Gamma(m - 1) / 2 \; SQR(m/2))$$

Da der Korrekturfaktor Cm immer kleiner als 1 ist, ist die Varianz von g(u) immer auch kleiner als die von g.

"When the experimental and control group sample size are equal and a pooled standard deviation is used, the g(u) can be shown to be the uniformly minimum variance unbiased estimator of, meaning that anay other unbiased estimator must have larger varince than g(u)" (HEDGES 1980, S. 27).

Die Anwendung der unterschiedlichen Standardisierungskoeffizienten findet in bisher durchgeführten Meta-Analysen ihren Niederschlag. Einen wichtigen Hinweis liefern KULIK, KULIK, & COHEN (1979). Sie fragen danach, wieweit die verschiedenen Methoden zur Bestimmung der Effektgrößen übereinstimmen. Exemplarisch berechnen sie neben der Effektgröße (sensu GLASS) den COHEN-Parameter 'd', Differenzen in Prozentwerten, sowie Effektschätzungen auf einer vierpoligen Ratingskala. Die nachfolgende Tabelle stellt die Interkorrelationen dar:

Tabelle 9: Interkorrelationen zwischen unterschiedlichen Effekt-Koeffizienten nach KULIK, KULIK & COHEN, 1979, S. 327

	Diff. in Prozent	COHEN 's d	GLASS' ES
COHEN's d	.94	-	-
GLASS' ES	.95	.99	-
vierpolige Skala	.86	.84	.83

Aufgrund der hohen Korrelationen ist es möglich, mit Hilfe der Regression auf der Basis eines Meßwertes einen anderen vorherzusagen. Dies erlaubt den Umgang mit fehlenden Datenpunkten.

Der Einfluß des Treatmens auf die Varianz

Die Berechnung der Effektgröße unterstellt, daß Treatment wirke zwar auf den Mittelwert, nicht jedoch auf die Varianz, bzw. hier nur unwesentlich. In einer umfangreichen Integrationsstudie auf nicht-statistischer Basis zum Einfluß der Schule auf die Schüler unterscheiden NEWCOMB & FELDMAN (1969) zwischen extensiver und intensiver Treatmentwirkung.

"It is clear that any observed freshman-senior difference is affected both by extensitiy and by intensitiy - that is, by te number of individuals who change, and by the degree to which each of them changes" (S. 53).

Daraus sind folgende Kombinationen ableitbar:

		INTENSIV: HOCH	INTENSIV: NIEDRIG
EXTENSIV	HOCH	relativ hohe Mittelwertdifferenzen bei 'gleichbleibender Varianz	relativ niedrige Mittelwertdifferenzen bei 'gleichbleibender Varianz
EXTENSIV	NIEDRIG	relativ niedrige Mittelwertdifferenzen bei hohem Varianzunterschied	relativ niedrige Mittelwertunterschiede bei hohen Varianzunterschieden

Beispiel:
Drei Studien untersuchen die Wirkung unterschiedlicher Unterrichtsmethoden bei unterstellter Bedingung, die Vortestdaten seien identisch und bilden die Kontrollgruppendaten ab. Jede Studie untersucht 10 Probanden pro Gruppe. Für die Probanden liegen die folgenden Werte vor:

Tabelle 10: Beispieldaten für extensive und intensive Treatmentwirkung

	Kontroll-gruppe*	Treatmentgruppen I	II	III	IV
		E=h I=h	E=h I=n	E=n I=h	E=n I=n
01	12	18	16	12	12
02	14	19	14	24	14
03	12	17	18	12	15
04	14	21	20	14	14
05	16	22	18	26	19
06	14	21	16	24	14
07	14	21	18	14	14
08	16	21	18	16	19
09	18	24	20	18	18
10	14	20	16	14	14
AM	14.4	20.4	17.4	17.4	15.3
SD	1.84	2.01	1.9	5.34	2.45
ES/K		3.26	1.63	1.63	.49
ES/p		3.05	1.60	.75	.42
ES/T		2.99	1.58	.56	.37

E = Extensität, I = Intensität, h = hoch, n =niedrig, ES/K Effektgröße auf der Grundlage der Standardabweichung der Kontrollgruppe, ES/p = Effektgröße auf der Grundlage der gepoolten Varianz, ES/T = Effektgröße auf der Grundlage der Stadardabweichung der Treatmentgruppe. * Für die Daten wird unterstellt, daß die Gruppen sich im Vortest nicht unterscheiden.

In diesem Fall suggerieren dieselben Effektgrößen denselben Effekt, obgleich die Rohdaten unterschiedliche Wirkungen widerspiegeln (Treatmentgruppe II und III). Die gepoolte Standardabweichung 'kompensiert' solche Unterschiede und wird in der Diskussion mit Verweis auf die Unempfindlichkeit gegenüber der Homogenitäts- und Normalitätsannahme (vgl. u.a. KERINGER, 1964; COHEN, 1969; WINER, 1969) beim t-Test auch dort angewandt, wo die Varianz sich aufgrund der Treatmenteinwirkung verändert (GLASS 1979, HALL 1978 , KULIK 1980).

Effektschätzungen aus anderen Statistiken

GLASS, McGAW & SMITH (1981) verweisen darauf, daß Experimentalvergleiche häufig komplexer als der Vergleich von zwei Gruppen ausfallen. Häufig stehen damit auch nur Statistiken für zusammengefaßte Dateninformationen zur Verfügung. Zur Schätzung der entsprechenden Effektgrößen liefern sie verschiedene Hinweise.

Liegt eine Varianzanalyse für zwei Faktoren vor (z.B. Faktor I: Treatmenteinfluß, Faktor II: Geschlechtseinfluß), läßt sich die Standardabweichung wie folgt schätzen:

$$s = \sqrt{\frac{SS_{II} + SS_{I,II} + SS_{W}}{df_{II} + df_{I,II} + df_{W}}}$$

Der Effekt männlicher Probanden läßt sich dann berechnen:

$$((AM(T,m) - AM(C,m)) / s'$$

wobei gilt:

$$s' = SQR(MS(w))$$

Fehlt es im Vergleich von zwei Treatments an einer echten Kontrollgruppe (z.B. Programm-Evaluation), kann die Effektgröße in Referenz zur Vergleichsgruppe kalkuliert werden:

Beispiel:

Programm I: neue Unterrichtstechnologie
Programm II: traditioneller Unterricht
Programm III: neue Unterrichtstechnologie

$$ES = (AM(PI) - AM(PII)) / SD(PII)$$

Der traditionelle Unterricht fungiert hier als 'Kontroll'-Gruppe. Werden mehrere Programme verglichen, kann die Überlegenheit zwischen Programm I und III als Differenz ausgedrückt werden. GLASS, McGAW & SMITH (1981) empfehlen, die Standardabweichung der jeweiligen Referenzgruppe dann zu schätzen, wenn Untersuchungen vorliegen, die die Programme II und III vergleichen:

$$s'(c) = bo + b1*s(P\ III)$$

Die Autoren verweisen auf die Möglichkeit, referenzbezogene Mittelwerte in ihre Bestandteile zu zerlegen, wenn verschiedene Studien die Programme I und II, andere Studien die Programme I und III und wieder andere die Programme II und III untersuchen.

Nicht selten fehlen sowohl Mittelwerte wie auch Standardabweichungen. Für Untersuchungen, in denen mehrere Experimentalbedingungen verglichen werden und Mittelwerte nicht angegeben sind, besteht kaum eine Möglichkeit, Effektgrößen zu rekonstruieren. Ein t-Wert wird für nicht-aggregierte Statistiken in die Effektgröße transformiert:

bei ungleichen n: $\hat{\Delta}_{T-C} = t\sqrt{1/n_T + 1/n_C}$

bei gleichen n: $\hat{\Delta}_{T-C} = t\sqrt{2/n}$

Empirische oder kritische Wahrscheinlichkeitswerte werden nicht über die z-Transformation sondern über die Transformation der t-Verteilung in Effektgrößen umgewandelt. Diese Schätzung ist nach Angaben der Autoren konservativer (vgl. S. 129f).

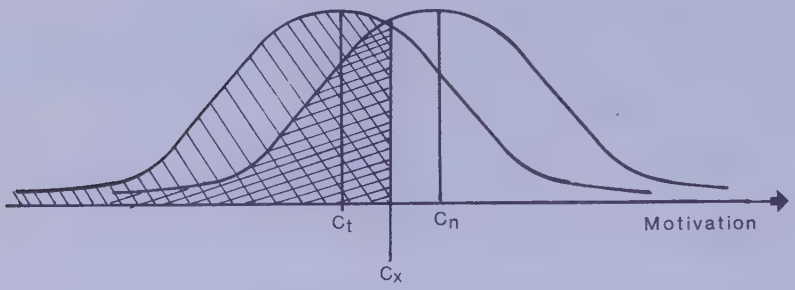
Für nicht-parametrische Statistiken empfehlen die Autoren ebenfalls die Transformation des entsprechenden Signifikanzwertes in den zugehörigen t-Wert. Dabei ist jedoch die niedrigere Teststärke der nicht-parametrischen Verfahren zu berücksichtigen.

Nicht selten sind Experimentalbefunde durch dichotome Variablen ausgedrückt (z.B. krank versus gesund, kooperativ versus nicht-kooperativ). Es ist unangebracht, aus diesen Daten unmittelbar eine Effektgröße berechnen zu wollen. GLASS, McGAW & SMITH (1981) empfehlen für solche Fälle, eine (unbeobachtete) metrische Information zu unterlegen. Angenommen wird, daß Kontroll- und Experimentalgruppen normalverteilt sind.

Beispiel nach GLASS, McGAW & SMITH
Die Wirkung eines neuen Curriculums soll u.a. an der Drop-Out-Rate von Schülern gemessen werden. Die abhängige Variable ist dichotom: Ausstieg versus Verbleib in der Schule. Als unbeobachtete Variable wird die Motivation, in der Schule zu bleiben, unterlegt. Fällt die Motivation unter einen kritischen Wert (cut off point C(x)), steigt ein Schüler aus der Schule aus. Bekannt sind die Prozentanteile der verbleibenden (bzw. aussteigenden) Schüler in Abhängigkeit des Curriculums:

traditionelles Curriculum	[C_t]	40% Nicht-Aussteiger	p(c)
neues Curriculum	[C_n]	60% Nicht-Aussteiger	p(e)

Abb. 3: Motivation als unterlegte Metrik für die Dropout-Rate in Abhängigkeit der Curricula (in Anlehnung an GLASS, McGAW, & SMITH (1982)



Die Wahrscheinlichkeiten, mit denen Schüler in Abhängigkeit des Curriculums aussteigen, können beobachtet werden. Wird für Motivation eine Normalverteilung unterstellt, gilt:

$$p(e) = \int_{-\infty}^{z(e)} \frac{1}{\sqrt{2\,pi}} \; e^{-z^2/2}\, dz$$

wobei

$$z(e) = \frac{X - AM(X(e))}{s(e)}$$

entspricht.

z(e) entspricht der Standardnormalverteilung, welche die Kurve an dem 100p(e)-ten Perzentil teilt. Entsprechendes gilt für p(c) bzw. z(c). Da nun

$$z(e) = (C(x) - AM(X(e))) \;/\; s(e)$$

und

$$z(c) = (C(x) - AM(X(c))) \;/\; s(e)$$

kann bei unterstellter Varianzhomogenität die Effektgröße bestimmt werden:

$$A = z(C) - z(T) = \frac{AM(T) - AM(C)}{s(x)}$$

Integration von Korrelationsstudien

Die Meta-Analyse ist auch auf Korrelationsstudien angewandt worden. WHITE (1976) untersucht den Zusammenhang zwischen sozioökonomischem Status und Schulleistung auf der Basis von 100 Studien mit insgesamt 620 Korrelationskoeffizienten.

"The mean magnitude of the correlation coefficient taken across all studies and collapsing over all others characteristics was .315 with standard deviation of .225. The median r was .251 indicating a positively skewd distribution..." (S. 76).

Ein besonderes Problem bei Korrelationskoeffizienten besteht in der Verwendung der Analyseeinheit. ROBINSON (1950) weist nach, daß Korrelationskoeffizienten auf der Basis von Aggregaten (z.B. Schulklassen, Schulen) wesentlich höher ausfallen als Korrelationskoeffizienten auf der Basis von Individuen. WHITE (1976) zeigt dieses Problem auf und berechnet die mittlere Korrelation in Abhängigkeit der jeweiligen Aggregate.

Korrelationsmittelungen sind nicht verzerrungfrei, da sie häufig auf der Basis der Aggregationen von Klumpen entstehen. Auf dieses Problem verweist HORNKE (1973).

"Durch die institutionell vorgegebenen Klassen können sich Effekte in das Datenmaterial einschleichen, die die Interpretation verfälschen" (S. 88).

Eine ungewichtete Mittelung ist nur dann angebracht, wenn die Stichproben sich hinsichtlich ihrer Statistiken bzw. die Populationen sich hinsichtlich ihrer Parameter nicht unterscheiden. Gegenüber der arithmetischen Mittelung fließen in das gewogene Mittel zumindest die unterschiedlichen Gruppenfrequenzen mit ein:

$$\overset{\circ}{r} = \sum \frac{n_k}{n} r_k$$

Einen Koeffizienten, der die Varianzunterschiede einbezieht, liefert der MLE-Koeffizient nach OLKIN (vgl. HORNKE 1973):

$$\overset{\circ}{r} = \frac{\rho \sum n_k}{1-\rho^2} = \sum \frac{n_k \; r_k}{1-r_k \rho}$$

GLASS 1978), McGAW, GLASS & SMITH (1981), WHITE (1976) liefern verschiedene Verfahren zur Transformtion unterschiedlicher Korrelationskoeffizienten.

1.4.6 Berechnung mittlerer Effekte und Varianzaufklärung

In Abhängigkeit der Fragestellungen werden die Effektgrößen gemittelt. Bei mehreren Effekten aus derselben Studie empfiehlt sich die Anwendung der Methode des Jackknifing (vgl. GLASS 1978). Die gemittelten Effektgrößen lassen sich sich graphisch in Bezug zur Kontrollgruppe darstellen (vgl. Abb. 2). Eine andere Darstellungsweise, die die Lage von Extremwerten wie die Verteilung der Werte in Abhängigkeit vom Median anzeigt, liefert die explorative Methode nach TUKEY (1977). DIXON & MASSEY (1969)verweisen auf die Möglichkeit, solche Extremwerte auf ihre Zugehörigkeit zur Population des Datensatzes zu überprüfen und ggf. durch Windsorisierung die Extremwerte zu egalisieren (Varianzhomogenisierung).

Mit Hilfe der multiplen Regressionsanalyse kann die Varianz der Effekte auf systematische Einflußgrößen untersucht werden. Systematische Einflusse können entweder aufgrund theoretischer Annahmen oder durch die Entdeckung von Unterschieden während der Meta-Anlyse einfließen:

ES' = Treatmenteffekt + Modertor-Effekte + Effekte durch Hintergrundtheorien + e

HEDGES (1982a, 1982b, 1982c) liefert eine generelle Strategie, die die Modellanpassung an die Effektgrößen in Abhängigkeit der Datenqualität (kategorial vs kontinuierlich) erlaubt.

1.4.7 Ausgewählte Probleme zur Meta-Analyse

Der Beitrag von GLASS (1976) hat nicht nur zu einer umfangreichen Publikation uber die Verfeinerung der meta-analytischen Integration und Beitrage zur Durchführung von Meta-Analysen hervorgerufen. Eine Vielzahl von Autoren stehen diesem Verfahren äußerst distanziert gegenüber. Die Kritik richtet sich insbesondere auf zwei Aspekte. Zum einen wird der Vorschlag von GLASS (1978) bemängelt, auch Studien einzubeziehen, die eine schwache methodische Validität ausweisen, zum anderen wird der Vorwurf erhoben, die Integration über viele Studien hinweg führe zu einer Konfundierung von Information und nicht zu deren Systematisierung.

Design-Qualität

GLASS (1976) , GLASS (1977), GLASS (1978) schlägt vor, Studien unabhängig von ihrer Design-Qualität in das Syntheseverfahren aufzunehmen. Damit bricht er die Konvention, nur 'gut geplante' Studien zum Ausgangspunkt von Schlußfolgerungen zu machen.

"Research design has a logic of its own right, but it is not a logic appropriate to research integration" (GLASS, 1978b, S. 355).

Studien mit schwachen Designs können dann als hinreichend valide betrachtet werden, wenn ihre Ergebnisse mit den Studien guter Designs übereinstimmen.

Beispiel I
Es liegen zwei Studien vor. Studie 1 ist gut geplant, Studie 2 liefert Vortest- und Nachtestergebnisse ohne Kontrollgruppe. Beide Untersuchungen kommen zu ähnlichen Ergebnissen. Wenn der Vortestwert eine gute Schätzung für den Nachtestwert ohne Treatmenteinfluß ist, die 'gute' Studie mögliche Störvariablen ausschließt, alternative Erklärungen für beide Studien durch den 'gesunden Menschenverstand' (vgl. COOK & CAMPBELL (1981)) ebenso ausgeschlossen werden können, gibt es keinen Grund, das Ergebnis der schlecht geplanten Studie nicht zu akzeptieren.

Beispiel II
Für eine Integration werden 100 Studien zusammengetragen. Sie beinhalten alle jeweils eine Schwäche (fehlende Repräsentativität, schlechtes Meßinstrument, schwache interne Validität usw), sind aber ansonsten gut geplant. Wenn alle Studien ähnliche Ergebnisse zugunsten eines Treatments in Relation zur Kontrollgruppe aufweisen, dann kann nach GLASS (1978b) das Überalles-Ergebnis nur mit dem Hinweis verworfen werden, die Treatmentwirkung wäre auf eine multiple Kausalität zurückzuführen, d.h. sie wäre durch fehlende Repräsentativität oder schlechte Meßinstrumente oder schwache interne Validität hervorgerufen.

Mit diesem Argument unterscheidet GLASS die Rationalität der Designbewertung von Einzelstudien gegenüber der von Integrationen. In der Planung einer Einzelstudie sind soviel alternative Variablen wie möglich auszuschließen. Bei der Integration entscheidet es sich jedoch a posteriori, ob die auf de Design-Qualität zurückführbare Varianz erklärenden Einfluß ausübt oder nicht. Dies ist allein eine empirische Frage, der entsprechend explorativ anhand der vorliegenden Daten nachgegangen werden kann. In ihrer Meta-Analyse zu Therapieeffekten konnten GLASS und Mitarbeiter zeigen, daß die Design-Qualität dort keine bedeutende Rolle spielt. Die Korrelation zwischen interner Validität und Effektumfang betrug r = - .09. Im Gegensatz hierzu fanden die Autoren in ihrer Meta-Anlyse zum Zusammenhang zwischen Klassengröße und Einstellung zur Schule, "that well designed studies produced quite different results from studies with minimal controls" (GLASS & SMITH, 1980, S. 429). Der Einfluß möglicher Störvariablen ist nicht unabhängig von der Fragestellung. Je mehr alternative Erklärungen die Forschungsfrage involviert, desto größer ist die Wahrscheinlichkeit, daß der Einfluß des Designs Varianz erzeugt.

Die Philosophie der a posteriori-Bewertung der Design-Qualität hat eine massive Kritik ausgelöst. EYSENCK (1978) unterstellt GLASS & SMITH die Verteidigung eines schwachen Forschungstandards: "...it would mark the beginning of a passage into the dark age of scientific psychology" (S.69). MANSFIELD & BUSSE (1977) weisen auf die eigene Forschungspraxis hin, nach der gut-geplante Trainingsprogramme weniger wahrscheinlich zu signifikanten Trainingseffekten führten. Schlecht-geplante Designs können zu unechten Effekten führen, da (a) bei kleinen Stichproben große Effekte eher durch Zufall entstehen und (b) experimentelle Programme mit 'expertise and commitment' der Implementatoren konfundiert sind, die bei kleiner Stichprobe noch stärker wirken. Weiter existiert nach Ansicht dieser Autoren ein Publikationsbias zugunsten gut-geplanter Studien unabhängig von ihren Ergebnissen, während schlechte Studien dann wahrscheinlicher veröffentlicht werden, wenn sie signifikante Ergebnisse aufweisen. Die Autoren empfehlen daher, die schlecht geplanten Studien auszuschließen, um die restlichen Untersuchungen nochmals in gut-geplante Studien und in solche mit methodischen Grenzen zu unterteilen. Zeigen alle Studien in dieselbe Richtung, könne die Schlußfolgerung tentativ akzeptiert werden. Allein der Effektumfang aus den gut-geplanten Studien soll als Grundlage zur groben Schätzung des verallgemeinerbaren Umfangs dienen. Dieser Differenzierung schließen sich GLASS und Mitarbeiter durchaus an. In ihren Studien zum Zusammenhang von Klassenstärke und Schulleistung (GLASS & SMITH (1979)) und in der Antwort auf eine Kritik an dieser Studie durch den 'Educational Research Service' betont GLASS (1980):

"Since it makes a difference how well controlled a study is, I'll go to the best controlled studies to get my best guess of the magnitude of relationship" (S. 243).

JACKSON (1980) weist auf die Problematik hin, daß die Berechnung der Kovarianz zwischen Design-Qualität und Ergebnissen nur dann nicht zu fehlerhaften Schlußfolgerungen führt, wenn eine ausreichende Anzahl von Studien guter Design-Qualität zur Verfügung stehe. Zudem gebe es keinen Anlaß für die Annahme, die Beziehung zwischen Design-Qualität und Ergebnissen verlaufe linear oder monoton. Die Kongruenz in den Ergebnissen liefert keinen Hinweis auf die Validität. Bei hoher Kongruenz und fehlender Evidenz eines strengen allgemeinen Störfaktors für die Validität aller oder der meisten Studien, besteht zwar eine suggestive, jedoch keine konklusive Evidenz dafür, daß eine jede methodologische Schwäche, die in einigen Studien vorliegt, wesentlich auf die Ergebnisse dieser Studien wirkt.

Konsens bei den Meta-Analytikern besteht sicherlich darin, auch weniger gut-geplante Studien in die Datensynthese einzubeziehen. Bei übereinstimmung der Ergebnisse, d.h. bei fehlender Kovariation zwischen Design-Qualität und Ergebnissen können höchstens logische Argumente gefunden werden, die auf eine Datenverzerrung infolge der Untersuchungsqualität hinweisen. Existiert eine solche Erklärung nicht, darf die tentative Evidenz unterstellt werden, die Qualität der Untersuchung habe keinen Einfluß auf die Ergebnisse. Unterscheiden sich jedoch die Effektgrößen in Abhängigkeit der Design-Qalität, und weisen die Ergebnisse in dieselbe Richtung, sollten die Studien gewichtet sowie durch die Analyse eines Subsets der Studien der Einflußumfang bestimmt werden (nesting-analysis). Aussagen über die Richtung der Befunde sind dann auf der Grundlage aller Studien möglich, Ausagen über den Effektumfang sollten auf der Basis der gut-geplanten Studien gemacht werden.

Das Inkommensurabilitätsargument

Jenseits methodischer Verfeinerung und der Diskussion um Verfahrensfragen lautet die massivste Kritik an statistischen Integrationsverfahren, sie seien schlichtweg überflüssig, da man unterschiedliche Studien nicht integrieren könne. EYSENCK (1978) , KAZDIN (1978), GALLO Jr. (1978), HILGARD (1978) verweisen auf die Egalisierung und Verdeckung bdeutsamer Unterschiede zwischen Studien mit der Folge falscher Schlußfolgerungen aus der Integration.

"It can't be assumed that the other sources of variation in a large number of related studies will be random and can be treated as a background of noise against which the influence of selected variables stands out clear" (HILGARD 1978, S. 395).

Ähnlich argumentiert GALLO, wenn er auf die gegenwärtige (amerikanische) Forschungspraxis verweist:

"Too many studies...of psychology...are conducted in a matter of a few short weeks or month, use dependent measures chosen for convienience and simplicity of measurement, and are undertaken in the prevailing publish or perish' atmosphere... As long as the philosphy of our institution encourages these masive production of quick and easy studies, neither meta-analysist nor any other type of 'analysis' will enable us to answer our profession's most pressing questions" (GALLO, 1978, S. 695).

Auch EYSENCK bemängelt in scharfer Form die Unvergleichbarkeit einzelner Untersuchungen:

"the notion that one can distill scientific knowledge from a compilation of studies mostly over poor design relying on subjective unvalidated and certainly unreliable clinical judgement, and dissimilar with respect to nearly all the vital parameters, dies hard" (EYSENCK 1978, S. 697).

Die Konsequenz liegt demnach nicht in der Berücksichtigung des angesammelten Wissens, sondern in der Durchführung einer eigenen neuen Untersuchung, die Kraft der persönlichen Autorität besondere Evidenz erhält (vgl. auch AUSUBEL, 1983).

In weniger grundsätlicher Ablehnung wird auf die Gefahr hingewiesen, daß durch die Vielzahl der Studien wesentliche Detail-Information verloren gehen könne (vgl PRESBY, 1978, LIGHT & SMITH, 1971). Die Frage ist nur, welche Information von Wichtigkeit ist. GLASS (1978b, S. 358) begegnet diesem notwendigen Anliegen ein wenig scholastisch.

"in fields where many things 'work', determining what works best is the primary task. It requires a libriarian temperament and a solid techniqual expertise".

1.4.8. Eine Übersicht zu durchgeführten Meta-Analysen zwischen 1976 und 1980

Mit Erscheinen des Aufsatzes von GLASS (1976) wurden bis 1980 insgesamt 57 Meta-Analysen durchgeführt. In 39 dieser Datensynthesen flossen 2593 Einzelstudien ein. 25 Meta-Analysen berechneten insgesamt 11.270 Effektgrößen. Im Durchschnitt wurden pro Studien 7 Effekte extrahiert. Die wenigen Daten zeigen die Dichte der Information an. Die Schwerpunkte der Analysen lagen in der Unterrichtsforschung (36 Meta-Analysen). Die Tabellen 11 und 12 zeigen die Rangfolgen der untersuchten Variablen:

Tabelle 11: Häufigkeit der untersuchten abhängigen Variablen in 57 Meta-Analysen

Abhängige Variable	N	Abhängige Vaiable	N
Schülerleistung	22	Testüberprüfung	4
Fertigkeiten, Wissen, Verhalten	13	Vergessen/Behalten	3
Einstellung	6	Kommunikations-fertigkeiten	2
affektive Variablen	5	Locus of Control	2
Psychische Gesundheit	5	Lernen	2
kognitive Variablen	4	Kreativität	2

Tabelle 12: Untersuchte unabhängige Variablen in 57 Meta-Analysen

AV	UV-Kategorie	UV-Konstrukte
SCHÜLERLEISTUNG/ERWARTUNG	Vorzeichen-Variablen(1)	Curricula für die Elementarstufe Ernährungsprogramme
	Kontext-Variablen(1)	Advance Organizer traditioneller vs. offener Klassenraum Sozioökonomischer Status der Schüler Sonder- vs. reguläre Klassen Motivation der Schüler häusliche Umwelt Massenmedien Peer-Einfluß Sozialpsychologisches Klassenklima Klassengröße
	Prozeß-Variablen(1)	Lehrerverhalten Lehrereinstellung Lehrererwartung Instruktionstechniken Tutorsysteme
	Therapievergleich	Trainingsprogramme Verhaltens- vs Non-Verhaltenstherapie
	Therapiewirksamkeit	Drogentherapie Alkoholtherapie Transzendentale Meditation Hyperaktivität
	Sonstiges	Psychotherapie vs. medikamentöse Therapie
	Validität Theorie Psychologie	differentielle Validität Theorienevaluation Forschererwartungen

(1) Die Variablenunterteilung im Bereich der Unterrichtsforschung wurde DUNKIN & BIDDLE (1974) entnommen.

2 ZUR KOOPERATIONSFORSCHUNG

2.1 Zum Gegenstandsbereich

Das Phänomen 'Kooperation' ist so alt, wie Menschen denken können. Die christliche Lebensgemeinschaft, der Staat Platons sind frühe Darstellungsformen kooperativer Lebensgemeinschaften. Ein breites Spektrum sozialer Interkationsformen und kategorischer Imperative sind mit der 'Essenz' von Kooperation verbunden. DURKHEIM zeichnet die verschiedenen Formen humaner Interdependenzen und ihre Bedeutung für das Überleben in seiner Schrift zur Entwicklung von der 'mechanischen' zur 'organischen' Solidarität nach (1898) und setzt sich mit Ursachen und Erscheinungen der Verletzung kooperativer Interdependenzen (Der Selbstmord, 1897) auseinander. NISBET (1968) zeigt auf, daß Kooperation universell gelebt wird: sowohl im Pflanzen- wie im Tierreich; bei den Menschen in der kleinsten sozialen Einheit sowie zwischen souveränen Staaten. Kooperation als 'sozialer Prozeß' ist nach ihm in fünf Typen differenzierbar.

Die 'autonome Kooperation' basiert auf dem Instinkthaushalt und dem Standort im ökosystem. Sie äußert sich in koordiniertem, gemeinsamen Handeln. 'Traditionelle' Kooperation wird durch soziale Normen reguliert. Wechselseitige Hilfestellungen sind dabei von besonderer Bedeutung. Den Prototyp traditioneller Kooperationen bildet die Familie. Verschiedene, ökonomisch definierbare Lebensgemeinschaften - z.B. die dörfliche Gemeinschaft, zwischen deren Mitgliedern Wettbewerb verboten ist und die gegenseitige Hilfe bei der Aussaat oder der Ernte überlebensnotwendig sind - unterliegen diesem Regelsystem. Bei der 'vertraglichen' Kooperation sind die Handlungsinhalte präzisiert, spezifiziert und durch legale Sanktionen geregelt. Modernen Gesellschaften ist diese Kooperationsform inhärent. 'Gerichtete' Kooperation findet sich in Großorganisationen (z.B. Militär, Gewerkschaften, Großunternehmen, Schulsystemen). Die hierarchisch geordneten kooperativen Interaktionen entspringen dem Bewußtsein eines gemeinsamen Ziels oder klarer Normen. Die 'spontane' Kooperation ist demgegenüber situativ und freiwillig. Sie basiert auf häufig wiederkehrenden Ereignissen, die, durch Sympathie getragen, optimiert werden sollen (Elterninitiativen, Spielgruppen usw.).

Die vielfältigen Aspekte und Konzepte von Kooperation machen deutlich, daß in Abhängigkeit der unterschiedlichen Erklärungsansätze es DIE Theorie der Kooperation nicht gibt. Daher wird der psychologischen Kooperationsforschung, die sich mit den Bedingungen, unter denen sich Menschen kooperativ verhalten, und mit den Auswirkungen von Kooperation auf andere Verhaltensweisen beschäftigt, Theorielosigkeit vorgeworfen.

"...there has been little in the way of explicit theorizing and virtually no experimental work with respect to the effects of cooperation and competition upon the social process..." (DEUTSCH 1949, S. 129).

Obgleich in den letzten 35 Jahren die Anzahl der Untersuchungen unter dem Stichwort Kooperation stark angestiegen ist, wird die konstatierte Theorie- und Konzeptlosigkeit dieses Forschungszweiges auch heute noch unterstellt:

"...formal theory had not played a very large role in shaping the research ...psychologists have not developped a parsimonious set of integrating principles (a theory in other words) which will effectively serve to organize ... the rapidly accumulating empirical evidence" (HARTUP, 1965).

Zudem wird selbst dem theoretischen Rahmenwerken der Kooperationsforschung Ekklektizismus vorgeworfen (vgl. PHILLIPS & DE VAULT 1957, BRYAN 1975, WILKE-BIRKENHAUER 1976), der durch methodologische und methodische Vielfalt in den Untersuchungen zu sozial-interaktiven Phänomenen Unterstützung findet. Abweichende Ergebnisse, Widersprüche und Unvereinbarkeiten haben zudem das induktive Suchen nach einflußausübenden Faktoren unterstützt, insbesondere dort, wo das Meßinstrument selbst eine elaborierte Theorie verlangt. In der primär feldtheoretisch orientierten Forschung mit Hilfe des Gefangenen-Dilemma-Spiels (Prisoner's Dilemma Game, in der Folge PDG) nahmen entgegen der Erwartungen und theoretischen Vermutungen die kooperativen Wahlen mit zunehmenden Spieldurchgängen ab (vgl. BIXENSTINE & WILSON 1963, RAPAPORT & CHAMMAH 1965, SWINGLE & COADY 1967). So führt die Frage "Why then, 'Is selfdefeating behavior a predominant result?'" (BRAVER & BARNETT 1976, S. 161) zu einer Suche erklärender Variablen mit der Folge umfangreicher Hilfshypothesen. SCHMITT & MARWELL (1971a,b) diskutieren einen Risikofaktor; GUMPERT, DEUTSCH & EPSTEIN (1969) zeigen auf, daß ein monitärer Anreizwert keine Erklärung liefert und bestätigen die vorliegenden Ergebnisse. BRAVER & BARNETT (1976) versuchen zu zeigen, daß der Tendenz über die Spieldurchgänge hinweg mit Hilfe des Modellernens entgegengewirkt werden kann. Nach wie vor ist jedoch die Frage ungeklärt, welche multiplen Kontingenzen oder Bedingungen das Spielverhalten über eine Serie von Durchgängen beeinflussen (vgl. VINACKE 1969).

Der fehlenden theoretischen Stringenz stehen viele nominale und operationale Definitionen von Kooperation gegenüber. Jenseits der PDG-Forschung entwickelt nahezu jede Untersuchung ihr eigenes Meßinstrument und bedient sich eines anderen Vor-Verständnisses von Kooperation als Bestandteil eines (pro-) sozialen und interaktiven Verhaltensspektrums. Auf der anderen Seite werden in den publizierten Untersuchungen nicht selten Referenztheorien oder andere Referenzuntersuchungen angegeben, die die Deduktion der zur Untersuchung anstehenden Variable(n) suggerieren, unabhängig von der Frage, ob diese notwendig, hinreichend oder 'nur' fördernd wirken, und ohne die Stringenz der untergelegten theoretischen Paradigmen weiter zu diskutieren.

Einen nicht unwesentlichen Beitrag zur Systematisierung des Forschungsbereichs liefern theoretisch orientierte Beiträge (vgl. CROTT 1971, 1979, DEUTSCH 1949, 1962, HAKE & OLVERA 1978, JERSILD 1960, RAPOPORT & CHAMMAH 1965, STAUB 1978, VINACKE 1969) und qualitative Übersichtsartikel (vgl. BRYAN 1975, COOK & STINGLE 1974, HANNAH 1970, JOHNSON & JOHNSON 1974, KLIPPSTEIN & KLIPPSTEIN 1978, PHILLIPS & DeVAULT 1957). Sie explizieren einen jeweiligen Ausschnitt aus den vielfältigen Forschungsbefunden. In Einklang mit der vorherrschenden Übersichts-Methode werden zu unterschiedlichen Variablen empirische Arbeiten zitiert, die die Bedeutsamkeit der angesprochenen Variablen belegen sollen. Jenseits der Widersprüchlichkeiten durch unterschiedliche Methoden, Stichproben und Meßinstrumente extrahieren diese Übersichtsartikel ein Fülle untereinander kaum verbundener Variablen. Ihnen mag ohne Zweifel varianzaufkärende Funktion zukommen. Nur hilft diese Vermutung in keiner Weise weiter, die tatsächlichen, durch Interdependenzen unterschiedlich zu gewichtender Einflußfaktoren auch zu kennen.

Die Methodologie des kritischen Rationalismus verlangt eine andere Vorgehensweise, mit deren Hilfe es möglich ist, die Vielfalt der Ergebnisse im Nachhinein zu erklären. Diese 'Logik der Forschung' präjudiziert ein dem Übersichtsartikel entgegengesetztes Konzept. Nicht die Aneinanderreihung wirkender Variablen sollte im Brennpunkt der Darstellung stehen, sondern die jeweiligen Theoriekerne, einschließlich die auf das Phänomen gerichteten Hypothesengürtel. Mit dieser Strategie kann empirische Evidenz im doppelten Sinne hergestellt werden: zum einen als weitere Bestätigung oder Falsifikation des Theoriekerns wie als Bestätigung dafür, daß die aus dem Theoriekern deduzierten Hypothesen das Phänomen zu erklären in der Lage sind.

In der Folge werden die Theorien, auf die sich in großem Umfang die empirischen Untersuchungen zur Kooperation berufen, dargestellt. Zuvor wird die Beziehung zwischen Kooperation und prosozialem Verhalten aufgezeigt und die Konstruktinhalte zum Kooperationsbegriff herausgearbeitet.

Exkurs I: Zum Verhältnis von Kooperation und Prosozialität

Der Mensch als 'homo sociologicus' DAHRENDORF (1958) meint nicht nur, den Menschen zum Gegenstand soziologischer Betrachtung erheben zu können, sondern unterstellt das Soziale als 'fait social' (DURKHEIM 1895). Zumindest nach DURKHEIM wären Verhaltensweisen im sozialen Kontext auf ein Ziel gerichtet bzw. auf ein gemeinsames Ziel unterschiedlicher sozialer Gruppierungen. Die Negation des Begriffes 'sozial' träfe höchsten den kontemplativ-vereinsamten Eremiten, nicht aber das wertend gemeinte unsoziale Verhalten. HINDE (1974) weist auf die Schwierigkeit funktionaler wie kausaler Klassifikationen von Verhalten.

"Thus definitions neither in causal nor in functional terms is wholly satisfactory, and it is better to return to common usage . Social behavior is a category of patterns of behavior that leads to the proximity of or involve interactions with, another individual of the species" (S. 15).

Dieser Annäherungsversuch ist nun nicht eindeutiger als soziologische Standortbestimmungen auch. Prosoziales Verhalten wird demgegenüber als Antithese zum wertend gemeintem aggressiven Verhalten kontrastiert (vgl. E.Z. JOHNSON 1951). Es beinhaltet die Hilfe gegenüber anderen Mitmenschen (vgl.STAUB 1979), die durch internale Normen, so die Annahme, reguliert wird (vgl. BRYAN 1975, ROSENHAN 1972). Subsummiernd wird unter prosozialem Verhalten Sympathie, Geben können, Hilfe, Kooperation, Unterstutzung, Schenken aufgezählt (vgl. WISPE 1972, STAUB 1979, BRYAN 1975). Als Prototyp prosozialen Verhaltens finden 'Altruismus' (vgl. CAMPBELL 1972, COHEN 1962, ROSENHAN 1972, KATZ 1970) und 'helfendes Verhalten' (vgl. BRYAN 1975, STAUB 1979, WALSTER & BERSCHEID, & WALSTER 1973, HORNSTEIN 1972) besondere Aufmerksamkeit in der Konfrontation mit Theorien (vgl auch HECKHAUSEN 1980). WISPE (1972) fragt nach der Logik verschiedener Klassifikationen und bietet eine Arbeitsdefinition prosozialen Verhaltens an:

"Behavior characterized as positive or prosocial, if generalized to most social situations, would be expected to produce or maintain the physical and psychological well-being and the integrity of the other person(s) involved. This kind of behavior intends not only the well being of the other person(s), but also a willingness to share, however briefly, not only the other's ends-in-view, but even his pain, frustration, and sorrow. This must be accomplished while the self clearly maintains its distinction from the other person(s).
By contrast, the kind of behavior defined as negative social behavior, if pursued, would circumscribe or destroy the freedom, expression, integrity, and physical or psychological well-being of the other person(s). Negative kinds of behavior are self-oriented, being concerned neither with the pleasure nor the welfare of the other person(s)" (S. 7).

Diese und auch andere Definitionen (z.B. STAUB 1979a) betonen den gebenden Aspekt prosozialer Verhaltensweisen. HAKE & VUKELICH (1972) differenzieren auf den Dimensionen 'gebender' und 'nehmender' Reaktionen verschiedene Formen prosozialer Verhaltensweisen. 'Maximale Kooperation' unterscheidet sich durch eine annähernde Gleichverteilung der gebenden Reaktionen von 'minimaler Kooperation'. Diese Gleichverteilung , von den Autoren als 'equity'-Prinzip der Kooperation eingeführt, entsteht durch eine hohe Korrespondenz gebender Reaktionen. Teilen (sharing) ist durch die Tendenz zur Gleichverteilung der Belohnung auf der Dimension nehmender Reaktionen ausgezeichnet, so z.B. wenn sich zwei Kinder Bonbons durch wechselseitige Entnahme dieser Süßigkeiten aus einer Tüte teilen. Kompetition unterliegt keinem equity-Effekt und besteht aus der Veräußerung nehmender Reaktionen. Altruismus bezeichnet demgegenüber ein Verhalten, das durch gebende Reaktionen (wieder) Gleichverteilung herstellt. Mit Hilfe der Reziprozitätsnorm

(GOULDNER 1961) lassen sich in Abhängigkeit des Gleichverteilungsprinzips folgende soziale Verhaltensweisen unterscheiden:

Abb. 4 Formen des Sozialverhaltens

Matrix 1
Sozialverhalten bei Gleichverteilung

	Person A: gebende Reaktionen	Person A: nehmende Reaktionen
R(G) / P(B)	maximale Kooperation	Altruismus von B gegenüber A
R(N)	Altruismus von A gegenüber B	Teilen

Matrix 2
Sozialverhalten bei vorliegender Ungleichheit

	Person A: gebende Reaktionen	Person A: nehmende Reaktionen
R(G) / P(B)	minimale Kooperation	Ausbeutung von B durch A
R(N)	Ausbeutung von A durch B	Kompetition

R(G) = gebende Reaktionen, R(N) = nehmende Reaktionen
P(B) = Person B

Exkurs II: Theoretische und nominale Definitionen von Kooperation

DEUTSCH (1949) liefert einen ersten Überblick damals vorherrschender Definitionen unter der Überschrift "Some existing formulations - A glance at the literature" (S. 130). Bis heute hat sich diese Betrachtung nicht verändert: "Disagreement primarily arises regarding the conditions necessary to characterize behavior as cooperative" (COOK & STINGLE, 1974, S. 918). HAKE & OLVERA (1978) äußern ihre Verwunderung, wie es möglich ist, die Vielfältigkeit der unter Kooperation beschriebenen Phänomene mit nur einem Begriff belegen zu wollen. Verschiedentlich wurde der Versuch unternommen, die operationalen Definitionen von Kooperation zu systematisieren (vgl. JOHNSON & JOHNSON 1975, COOK & STINGLE 1974, HAKE & VUKELICH 1972). Danach kann zumindest zwischen (a) funktional-analytischen (behavioralen) und (b) zielabhängigen Definitionen unterschieden werden. Residual verbleibt eine kleine Gruppe (c) minimaler Definitionen.

(a) Funktional-analytische Definition

Die Kontingenz, d.h. die Regel, die die Konsequenz auf ein Verhalten spezifiziert, erheben WEINGARTNER & MECHNER (1966) zum Klassifikationsrational sozialer Interaktion. Die einfachste Kooperation einer Dyade beschreiben die Autoren wie folgt:

"Here A's response produces a condition such that B's response will produce a positive reinforcement for both. The information is contained in the first line of the diagram. The second line indicates that those responses in the opposite sequence will have the same consequence" (S. 456).

ROSENBERG (1960) erweitert dieses Schema und nennt vier Merkmale, die eine kooperative Situation bestimmen:

(i) "There exist at least two outcomes for each individual following a response

(ii) At least one of these outcomes is rewarding (i.e., it provides positive reinforcement and/or terminates noxious stimulation) and at least one outcome is not rewarding (or less rewarding)

(iii) reward delivered to one person must not preclude its availability for the other

(iv) The outcomes to both persons are in part, or intirely a function of the response of the other person" (S. 318)

HAKE und Mitarbeiter (HAKE & VUKELICH 1972, HAKE & OLVERA 1978) präzisieren den funktional-analytischen Ansatz. Eine kooperative Reaktion unterliegt denselben Reaktionskontingenzen wie nicht-soziale Reaktionen. Der wesentliche Unterschied besteht darin, daß die Verstärkung an eine andere Person als die aktuell reagierende geht. Zwar kann auch die agierende Person eine Verstärkung erhalten, aber das Definitionsmerkmal besteht darin, daß den Verstärker der Interaktionspartner erhält (vgl. HAKE & OLVERA, 1978, S. 210). Kooperative Reaktionen werden dabei als gebende Reaktionen (R(G)) bezeichnet gegenüber solchen, die ausschließlich auf die eigene Bekräftigung abzielen. Sie sind als nehmende Reaktionen (R(N)) definiert:

$$(A)R(G) \longrightarrow (B)\ K(+) \quad \text{gebende Reaktion}$$

$$(A)R(N) \longrightarrow (A)\ K(+) \quad \text{nehmende Reaktion}$$

Ein kooperativer Effekt besteht aus einem Anstieg der gebenden Reaktionen. Dies bezeichnen HAKE und Mitarbeiter als minimale Kooperation. Die Kontingenz ist kooperativ, da der Verstärker zumindest teilweise von der Reaktion des Interaktionspartners abhängt:

$$(A)R(G) \longrightarrow (B)R \longrightarrow (B)\ K(+) \quad (\ \&\ (A)\ K(+))$$

Da nun der die gebende Reaktion äußernde Teilnehmer ebenfalls in den Genuß von Verstärkern gelangen möchte, muß sich der Ko-Akteur in der weiteren Folge kooperativ verhalten. Die Kooperative Kontingenz muß Reziprozität aufweisen:

$$(A)R(G) \longrightarrow (B)R \longrightarrow (B)\ K(+) \quad \text{I}$$

$$(B)R(G) \longrightarrow (A)R \longrightarrow (A)\ K(+) \quad \text{II}$$

Hohe Korrespondenz zwischen den gebenden Reaktionen beider Partizipanden führt zu maximaler Kooperation. (I) bezeichnet eine Kooperative Episode sowie (I + II) eine kooperative Periode.

(b) Zielabhängie Definitionen

Die häufig zitierte und zugleich elaborierteste Definition liefert DEUTSCH (1949). Das feldtheoretische Konstrukt der Zielregion steht dabei im Mittelpunkt:

"In a co-operative social situation the goals for the individuals or sub-units in the situations under consideration have the following characteristics: the goal region for each of the individuals or sub-units in the situation are defined so that a goal-region can be entered by any given individual or sub-unit only if all the individuals or sub-units under consideration can also enter their respective goal-region" (S. 131f).

Eine Person muß nach dieser Definition wahrnehmen, daß ihr Ziel in fördernder Weise vom Ziel der anderen Person abhängt. Interpersonal besteht dann eine kooperative Situation, wenn beide Personen - im Fall einer Dyade - ihre, sie fördernden Interdependenzen wahrnehmen (vgl. DEUTSCH 1962).

Bei HOMANS (1950) definiert die größere gemeinsame Gesamtbelohnung die Kooperation. Personen treten demnach dann in eine kooperative Interaktion, wenn die Gesamtbelohnung infolge ihrer Tätigkeiten größer ausfällt, als unter der Bedingung der individuellen Zielerreichung. Dies setzt nach NISBET (1968) ein gemeinsames Interesse oder die ähnlich erwartete Hoffnung auf Belohnung voraus. Zusammenfassend liefert PAULSON (1974) für die inhaltlich Gestaltung und Evaluation der SESAME STREET eine Definition, die diese Aspekte integriert:

"... cooperation refers to situations that actively involves two or more individuals in interaction. This interaction is instrumental: the partipants achieve some goals. It produces mutual benefit: each participant achieves a goal..." (S. 230).

Die austauschtheoretische Definition versucht, die ziel-interdependente Orientierung und den funktional-analytischen Ansatz zu integrieren. Mit Referenz auf AZRIN & LINDSLEY (1956), DANIELS (1942) und DEUTSCH (1949) liefern THIBAUT & KELLEY (1959) eine Definition der Kooperation. Voluntaristische soziale Interaktionen werden eingegangen, wenn Mitglieder in einer Interaktion perzipieren, daß die von ihnen zu erzielenden Ergebnisse infolge einer Kosten-Nutzen-Betrachtung oberhalb eines kritischen Entscheidungswertes liegen. Kooperative Interaktionen sind gegeben,

"if both persons are able to produce their maximum reward for the other at minimum costs to themselves. The relationship will not only provide each with excellent reward-cost position at the same time" (S. 31).

(c) Minimale Definitionen

Einige Untersuchungen liefern keinen Hinweis auf ein theoretisches Konstrukt. MEAD (1937) bezeichnet Kooperation als "act of working together to one end" (zit. nach DEUTSCH 1949, S. 130). EIFERMAN (1970) reduziert Kooperation auf die gemeinsame Zielperspektive:

"Still, cooperativeness is retained to an as yet undetermined degree, if each party contains more than one member with the members of each party sharing the same aim " (S. 580).

NELSON & MADSEN (1968) differenzieren zwischen Koordination und Kooperation.

"Cooperation is thought to be reflected when individuals coordinate their action in order to obtain what they want" (zit. nach BRYAN 1975, S. 130).

Zum aktuellen Stand der Kooperationsforschung

Forschungsbeiträge, die ausdrücklich - jedoch unabhängig davon, ob sie auch tatsächlich und in welchem Umfang - Kooperation untersuchen, lassen sich zwei Teilgruppen zuordnen: (1) in solche Untersuchungen, die Kooperation als unabhängige Variable und (2) in solche Untersuchungen, die Kooperation als abhängige Variable studieren. Zu beiden Bereichen liegen viele Einzelstudien und qualitative Übersichtartikel vor. In jüngster Zeit aggregierten JOHNSON et al. (1981) und SLAVIN (1980) mit Hilfe formaler Datensynthese die Ergebnisse der Kooperationsforschung, wobei Kooperation die unabhängige Variable für Leistungsverhalten und Gruppenkohäsion bildete. Auf der Basis von 122 Studien extrahierten JOHNSON et al. (1981) sowohl Effektgrößen, Vote-Counts und z-transformierte p-Werte. Als unabhängie Variablen verglichen sie Kooperation, Kooperation bei paralleler Gruppen-Kompetition, Kompetition und Individualverhalten. Die Autoren konnten zeigen, daß kooperative Gruppenstrukturen zu höherer Leistung führen als andere 'Motivationsstrukturen'. Neben dem Einfluß verschiedener Moderatorvariablen weisen die Regressionsanalysen auf drei wichtige Faktoren hin: (a) die Art der Aufgabenstruktur, d.h. ob die Aufgabe auch allein oder nur interdependent zu lösen ist, (b) der Aufgabentyp und (c) die Verstärkungsstruktur. Damit bestätigt diese Analyse die ein Jahr zuvor von SLAVIN (1980) extrahierten Einflußfaktoren. Kooperatives Lernen involviert demnach Verstärkungsinterdependenz, also eine Verstärkungsstruktur, die ungeachtet eines Referenzwertes hohe Individual- und/oder Gruppenleistung auch entsprechend hoch bewertet. SLAVIN (1980) benennt für den Schulbereich darüber hinaus individuelle Verantwortlichkeit, lehrerbezogene Struktuiertheit und die Anwendung kompetitiver Merkmale als wesentliche Einflußgrößen auf die Leistung durch kooperative Interaktionen. McGLYNN (1982) weist darauf hin, daß die Unterschiede in den Effektgrößen durch die verschiedenen Zielstrukturen auch auf Mediatorvariablen zurückgeführt werden können.

"An educational or industrial practitioner is always confronted with a specific set of circumstances" (S. 185).

Es reicht nach GLYNN (1982) nicht aus, allein aufzuzeigen, daß eine kooperative Zielstruktur hohe oder höhere Leistungswerte erreicht, es muß vielmehr aufgezeigt werden können, was eine kooperative Struktur auszeichnet

und wodurch sie aufrechterhalten wird. Diese 'Aufklärungsarbeit' leisten Studien, die danach fragen, unter welchen Bedingungen sich Individuen kooperativ verhalten. Die vorliegende Meta-Analyse soll hierzu die wesentlichen Bedingungen nennen. Der nachfolgende Theorieteil entstand aus dem Vorwurf der Theorielosigkeit und des Ekklektizismus innerhalb der Kooperationsforschung; wie verschiedene Hinweise bereits anzeigen, findet sie nicht ganz ohne Theorienbezug statt.

2.2 Theorien zur Erklärung kooperativen Verhaltens

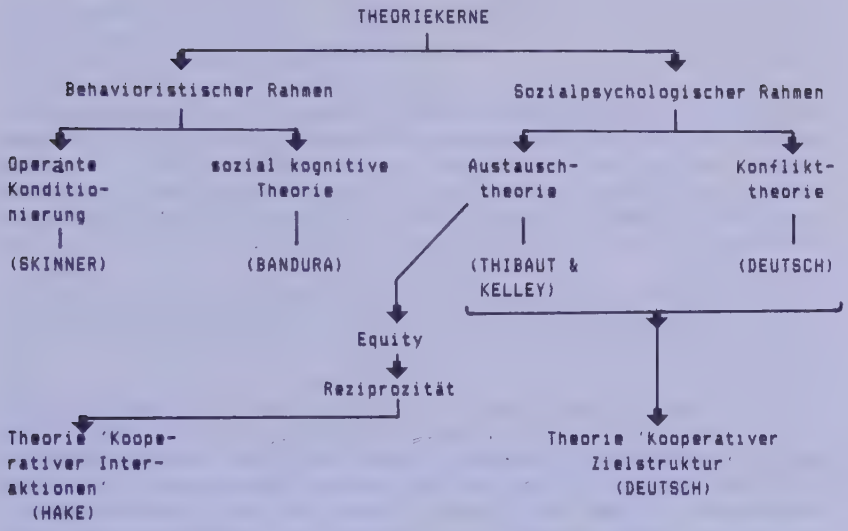
Die theoretischen Ansätze zur Erklärung prosozialen Verhaltens sind verschiedentlich klassifiziert worden. NACSON (1976) unterscheidet zwischen 'Reinforcementtheorien', 'mathematischen Modellen', 'gruppendynamischen Ansätzen' und 'sozialisationstheoretischen Ansätzen'. BRION-MEYSEL (1977) differenziert zwischen einem 'nature approach', in dem die Betonung auf biologische Variablen liegt (FREUD, LORENZ, JENSSEN), dem 'nurture approach', der das Sozialverhalten auf Umweltkontingenzen zurückführt und dem 'interactionist approach', in dem die Interaktionen zwischen verdeckten und offenen Variablen betont werden und mit Namen der kognitiven Entwicklungstheoretiker (KOHLBERG, PIAGET, SELMAN) verbunden sind. Die folgenden Theoriekerne werden genannt:

(1) die sozial-kognitive Entwicklungstheorie, einschließlich eines entwicklungstheoretischen Ansatzes, der ohne Verweis auf eine theoretische Referenz die Beziehung zwischen Kooperation, so wie sie das Meßinstrument definiert, und dem Lebensalter untersucht.

(2) Der lerntheoretische Ansatz, der sowohl die Theorie der operanten Konditionierung wie die sozial-kognitive Theorie des Lernens nach BANDURA involviert. Hier wird als abhängige Variable sowohl Verhalten im natürlichen Umfeld, an Hand eines handlungsbezogenen Meßinstruments wie an Wahlverhalten (z.B. PDG) untersucht. Hierunter wird auch das von HAKE und Mitarbeitern konzipierte kooperative Verhalten unter Einbeziehung austauschtheoretischer Paradigmen subsummiert. Die austauschtheoretischen Ansätze werden anhand der Austauschtheorie von THIBAUT & KELLEY (1959) sowie der Equity-Theorie von WALSTER und Mitarbeitern (vgl. WALSTER, E. & WALSTER, G.W., & BERSCHEID, E., 1978) dargestellt.

(3) Einen letzten Ansatz, der seinerseits auf austauschtheoretische Aspekte zurückgreift und sich in seinem Kern aus der Feldtheorie LEWIN's (1939) entwickelte, bildet die Kooperationstheorie von DEUTSCH (1949). Abschließend werden die Kooperationsannahmen aus theoretischen Ansätzen vom Meßinstrument her dargestellt. Die Untersuchungen im Rahmen der Nicht-Nullsummen-Spiele (PDG) bilden nahezu eine Untersuchungstradition für sich.

Abb. 5: Theoriekerne in der Übersicht



Da in die vorliegende Meta-Analyse keine Studien einflossen, die über die Alters-Variable hinaus Hypothesen im Licht der sozial-kognitiven Theorien nach PIAGET und KOHLBERG überprüfen, wird diese Theorie im der Folge nicht dargestellt. Ebenso orientiert sich die Beschreibung der Theorie von DEUTSCH im wesentlichen an seinem Ursprung, der Feldtheorie und weniger an seiner Elaboration, der Konflikttheorie.

2.2.1 Lerntheoretische Erklärungen

Operante Konditionierung

Der als Anti-Theoretiker etikettierte SKINNER (vgl. SKINNER 1969) liefert eine Theorie des Verhaltens, dessen unabhänige Variablen auf physikalische Elemente, nämlich Reize aus der Umwelt, zurückzuführen sind. Der Mensch verhält sich in Folge seiner Konsequenzen (Verstärkungsparadigma), die durch - dem Verhalten vorangestellte - diskriminative Stimuli angekündigt werden (Stimulsparadigma). Damit verzichtet SKINNER auf jede Form 'mentaler' Variablen. MAHONEY (1974) spricht hier in Abgrenzung kognitiver Verhaltstheorien von einem nicht-vermittelndem S-R-Modell. Die Ablehnung intervenierender Variablen bestimmt zugleich das methodische Konzept von

SKINNER. Der wissenschaftstheoretischen Tradition BRIDGMAN's (1927) und MACH's (1911) folgend, schreibt SKINNER (1974, englisch 1969):

"Verhalten ist einer der Gegenstände, die ohne hypothetisch-deduktive Methoden auskommt. Sowohl das Verhalten selbst, als auch die meisten Variablen, von denen es eine Funktion ist, sind gewöhlich ohne weiteres einsehbar" (S.10).

Wissenschaft (als eine Form des Verhaltens) kommt ohne hypothetisch-deduktive Methoden aus. SKINNER selbst verzichtet nicht ganz auf hypothetische Elemente, auf die sein funktionales Verhaltens-Modell zurückgeführt werden kann. Aus der Biologie übernimmt er die Equilibriums-These: der Mensch strebt nach Wohlbefinden des Organismus und zeigt dabei 'biologisch signifikantes Verhalten'.

"Doch können wir glaubhaft argumentieren, daß die Fähigkeit, sich durch ein Feedback aus der Umwelt verstärken zu lassen, biologisch vorteilhaft sein wird, da dieses den Organismus darauf vorbereiten soll, auf die Umwelt erfolgreich einzuwirken (SKINNER, 1973, englisch 1953, S. 86).

Nur so, schließt SKINNER, bietet dieser Erklärungsansatz für die funktionale Analyse des Verhaltens keine Hilfe. Sein quasi-kausaler Ansatz baut vielmehr auf die Ergebnisse insbesondere von PAVLOV (1927) und THORNDIKE (1911) auf. Mit ihnen ist das Paradigma von der klassischen Konditionierung verbunden. Wenn ein neutraler Reiz mit einem einen Reflex auslösenden Reiz zeitlich und häufig gekoppelt ist, übernimmt der anfänglich neutrale Reiz die (Verhaltens-)auslösende Wirkung:

UCS = unkonditionierter Stimuls, UCR = unkonditionierte Reaktion, NR = neutraler Reiz, CS = konditionierter Stimulus, CR = konditionierte Reaktion.

Die Pavlov'sche Konditionierung eines Hundes ist hierfür ein bekanntes Beispiel. THORNDIKE (1898) zeigt an seinem Katzenexperiment anhand von 'Lernkurven' die Wirkung von Verhaltenskonsequenzen. Dieses Effektgesetz (law of effect) verweist auf die Steuerung von Verhalten nicht nur durch vorausgehende Reize, sondern auch aufgrund nachfolgener Stimuli. Diesen Zusammenhang greift SKINNER in seiner Funktionalen Analyse auf. Er betont den - auf den ersten Blick einfachen - Mechanismus, nach dem ein Verhalten, welches positive Konsequenzen nach sich zieht, in der Zukunft häufiger und/oder wahrscheinlicher (in ähnlichen Situationen) auftreten wird:

$$p\ (R) / K(+) \longrightarrow P\ (R)$$

p ist P

Eine, durch nachfolgende Reize bzw. Konsequenzen kontrollierte Verhaltensform bezeichnet SKINNER als Operanten:

"Der Operant wird definiert durch die Eigenschaft von der oder auf die die Verstärkung abhängig oder kontrolliert ist" (SKINNER 1973, S. 70, englisch 1953).

Mit einem Operanten wird weniger ein einzelner Verhaltensakt als vielmehr eine 'zusammengehörige Reihe von Akten' bezeichnet. Positive Konsequenzen heißen positive Verstärker. Entweder existieren sie in Form natürlicher (primärer) Verstärker oder in der Form der durch den Mechanismus der klassischen Konditionierung erzeugten konditionierten (sekundären) Verstärker. JEHLE (1978) unterscheidet zwischen materiellen (einschließlich symbolischen), sozialen und Aktivitätsverstärkern (S. 75). Ein generalisierter Verstärker liegt vor, wenn er mit mehr als mit einem primären Verstärker gekoppelt ist:

"Eine Art generalisierter Verstärker wird dadurch erzeugt, daß viele primäre Verstärker erst dann richtig angenommen werden, wenn die physikalisch erfaßbare Umwelt bereits wirksam manipuliert worden ist" (SKINNER, 1973, S. 80, englisch 1953).

Ein Lebewesen wird nun nicht kontinuierlich verstärkt und verhält sich dennoch. SKINNER kann aufzeigen, daß intermittierende Verstärkung ausreicht, um Verhalten aufrecht zu erhalten. Partielle Verstärkungsbedingungen differenziert SKINNER nach fixen bzw. variablen Quoten. Diese Verstärkungsformen entscheiden neben der Latenzzeit über die Verstärkungseffekte, die Lerngeschwindigkeiten bzw. Löschungsresistenz (vgl. HOLLAND & SKINNER, 1961). Das Konzept der positiven Verstärkung wird um das der Strafe erweitert. Damit kann zwischen vier unterschiedlichen Konsequenzmöglichkeiten differenziert werden:

Abb. 6: Konsequenzmöglichkeiten und Folgen im Verhalten nach den Paradigmen der operanten Konditionierung

	K (+)	K (-)
gegeben	positive Verstärkung P(R) gt p(R)	Strafe P(R) lt p(R)
entzogen	Strafe P(R) lt p(R)	negative Verstärkung P(R) gt p(R)

Sowohl die nachfolgende Vergabe eines positiven wie auch die Zurücknahme eines aversiven Reizes wirken auf den Operanten verhaltenserhöhend. Die Wirksamkeit der Strafe als Instrument der Verhaltensformung (und der Erziehung) wird von den Vertretern der operanten Konditionierung stark in Frage gestellt (vgl. SKINNER 1969, HOLLAND & SKINNER 1961).

Die Wirkung von Verstärkungskontingenzen auf prosoziales Verhalten fassen HARTUP (1966), BRYAN & LONDON (1970) zusammen. Kritisch wird dabei angemerkt, daß die Ergebnisse auf die externe Gültigkeit durch Untersuchungen unter Laborbedingungen (HARTUP, 1965) und aufgrund der Einfachheit der Aufgaben bzw. Tätigkeiten (BRYAN & LONDON, 1970) beschränkt sind. In einer Übersicht liefert STEVENSON (1965) verschiedene, insbesondere die soziale Verstärkung beeinflussenden Variablen. So muß die experimentelle Aufgabe folgende Merkmale aufweisen, um nicht mit dem nachzuweisenden Effekt und der sozialen Verstärkung zu konfundieren: (i) die Aufgabe darf keine hohe intrinsische Motivation freisetzen, (ii) die Aufgabe darf keine klar erkennbaren Produkte liefern, (iii) die Aufgabe sollte die Ergebnisse früheren Lernens nur marginal berühren, (iv) die Aufgabe sollte dem Erwachsenen eine beliebige Kommentierung erlauben, (vi) sie sollte diskrete Reaktionen ermöglichen und (vi) reliabel sein. Nach STEVENSON (1965) haben 'marble dropping tasks' solche Merkmale.

GEWIRTZ & BAER (1958) wie auch STEVENSON (1965) zeigen auf, daß zwischen Verstärkungsrezipienten und Verstärkungsgebern Geschlechtsinteraktionen wirken. Bei weiblichen Experimentatoren wie z.B. Müttern, zeigen Jungen eine höhere Verhaltensrate, während bei männlichen Experimentatoren und Vätern Mädchen eine höhere Verhaltensrate äußern. Dies gilt nicht nur für das späte Vorschulalter, sondern auch für Schüler. STEVENSON (1965) weist jedoch darauf hin, daß männliche Verstärkungsinstanzen, die bei jungen

Kindern wirken, auch bei älteren Kindern wirksam sind, während weibliche Verstärkungsinstanzen, die bei jungen Kindern wirken, weniger Wirkung bei älteren Kindern zeigen (S. 103).

SKINNER (1969) betont, daß eine operante Konditionierung ohne Erwähnung des vorhergehenden Reizes beschrieben werden könne. Jedoch besteht Verhalten nicht ohne Relation zur Umwelt, so daß Reaktionen 'Anlässe' vorausgehen. Die hinweisenden Stimuli kundigen für die bestimmte Reaktion die eventuellen Folgen an:

$$S(D+) \longrightarrow R\ (p1) \longrightarrow K(+) \longrightarrow R\ (p2)$$

$$p(1)\ lt\ p(2)$$

Ihre hinweisende Funktion erhalten diese Stimuli ebenfalls nach den Prinzipien der klassischen Konditionierung, d.h. in der zeitlichen Assoziation eines bereits diskriminierten Reizes mit einem neutralen Reiz.

Die Organismus-Variable

Verhalten ist nicht allein eine Bedingung vorausgehender und nachfolgender Verstärkungskontingenzen. Die Begriffe 'Deprivation', 'biologisch-signifikantes' Verhalten weisen darauf hin, daß der Organismus selbst bei SKINNER kein 'horror vacue' ist. Jedoch besitzt im Rahmen dieser Theorie die Welt in unserer Haut' (SKINNER 1973, S. 239, engl.1953) für die funktionale Analyse des Verhaltens keine besonderen Eigenschaften. Diese Welt kann vielmehr mit Hilfe derselben Funktionsterme erklärt werden. MAHONEY (1974) verweist auf die Ironie, daß der als Vertreter der 'black box' geltende SKINNER eigentlich massiv ein vermittelndes Modell vertritt, d.h. ein Modell, das den Körper nicht ausspart.

"Obwohl er kein formalisiertes Vermittlungsmodell vorstellt, verteidigt er wiederholt die Annahme, daß die Prinzipien und Methoden der operanten Konditionierung auf Phänomene innerhalb der Organismen angewandt werden können" (MAHONEY, (dtsch) 1979, S. 73).

Verhalten im sozialen Kontext

Jede soziale Interaktion kann mit Hilfe der beschriebenen Variablen analysiert werden.

"Soziales Verhalten läßt sich als das Verhalten zweier oder mehrer Personen untereinander oder als deren gemeinsames Verhalten gegenüber einer gemeinsamen Umwelt definieren" (SKINNER, 1973, S. 275, engl. 1953)).

Die Analyse einer sozialen Interaktion (oder sozialen Episode) bezieht dabei das Verhalten der Akteure als determinierendes Verhalten ein.

"Wir können eine soziale Episode analysieren, indem wir uns mit jeweils einem Organismus zu einer gegebenen Zeit befassen. Unter den Variablen, die es zu berücksichtigen gilt, befinden sich die, welche ein weiterer Organismus hervorbringt. Dann setzen wir uns mit dem Verhalten des zweiten Organismus auseinander, wobei wir den ersten als Erzeuger von Variablen behandeln. Indem wir nun beide Analysen zusammenfügen, rekonstruieren wir die Episode" (SKINNER, 1973, S. 281, engl. 1953).

SKINNER selbst verweist auf Einflüsse, die die soziale Episode fördern oder behindern können. So entscheidet die Verfügbarkeit über Verstärker seitens der Akteure sowie die Reziprozität dieser Verstärkung die soziale Episode mit.

"Manchmal erklärt eine reziproke Interaktion das Verhalten anhand der Verstärkung. Jede Person hat ....etwas anzubieten, und sowie die Interaktion einmal zustande gekommen ist, erhält sie sich von selbst" (SKINNER, 1973, S. 286, engl. 1953).

Durch welche Mechanismen diese 'Autokatalyse' sozialer Interaktionen aufrechterhalten wird, erläutert SKINNER nicht. Die soziale Interaktion kann durch gemeinsame externe Stimuli beeinflußt werden. Mit Hinweis auf die Konkurrenzsituation , die sich durch eine Verstärkung des einen auf Kosten der Verstärkung des anderen auszeichnet, arbeitet SKINNER als bedeutende Merkmale der sozialen Interaktion die gegenseitige Verstärkung heraus. Konkurrierendes Verhalten muß daher nach SKINNER (1973, S. 287, engl. 1953) nicht unbedingt soziales Verhalen involvieren. Konkurrenz und Kooperation bilden nach seiner Ansicht nicht die Endpunkte eines gemeinsamen Kontinuums.

"Eine Kooperation, bei der Verstärkung zweier oder mehrer Einzelpersonen vom Verhalten dieser beiden oder aller Personen abhängt, ist offensichtlich nicht das Gegenteil von Konkurrenz, da sie ein verzahntes System zu erfordern scheint" (S. 87).

Neben der Reziprozität der Verstärkung sind die Mechanismen der personalen wie gruppengesteuerten Verhaltenkontrolle in sozialen Situationen wichtig. Kann Person A Person B vorsätzlich kontrollieren, ohne daß im gleichen Maße Person B die Person A zu kontrollieren in der Lage ist, spricht SKINNER von einer asymmetrischen sozialen Beziehung. Dieselbe Funktion mit erhöhter Wirkung kann eine Gruppe ausüben (group control). Hier deutet SKINNER (1973) auf einen wesentlichen Mechanismus hin, der für die Erklärung sozialer Interaktionen einbezogen werden sollte:

"Das Verhalten einer Person wird gewöhnlich dann als gut oder richtig bezeichnet, wenn es andere Gruppenmitglieder verstärkt, und als schlecht oder falsch, wenn es für sie aversiv ist" (S. 298).

Ein Verhalten, das den weiteren Gruppenmitgliedern eine Verstärkung in Aussicht stellt, wird zum diskriminativen Stimulus für die Reziprozität wie auch zum verstärkenden Stimulus selbst.

Die sozial-kognitive Lerntheorie BANDURAS

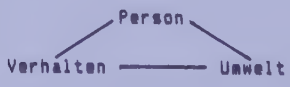
Die verschiedenen Theorien zum Modellernen wurden wiederholt klassifiziert. KANFER & PHILLIPS (1970) unterscheiden fünf Paradigmen des Modellernens. Als Sonderfall der operanten Konditionierung betrachten sowohl SKINNER (1953) wie auch GEWIRTZ (1969) den Prozeß des Beobachtungslernens. Theoretische Erklärungsversuche des Imitationslernens bestehen seit Ende des vergangenen Jahrhunderts. Als Kognitivisten werden häufig PIAGET, KOHLBERG, ARONFREED und BANDURA im Zusammenhang genannt. STAUB (1979) verweist jedoch auf den wesentlichen Unterschied zwischen kognitiven Entwicklungstheorien und der sozial-kognitiven Theorie des Modellernens:

"Development is used to refer to change with age due to maturation of varied competencies of the organism and/or to the unfolding of innate potetials in the course of experience, such as cognitive capacities or moral reasoning. As used by cognitive developmental psychologists, the term development also suggests that change is qualitiative...Learning theorists, in contrast, usually regard change as continuous,and what is learned is not considered as representing the unfolding of innate potentials but the acquisition of something by the organism" (S. 60).

Nach BANDURA (1973, dtsch 1979) stellt:

"Lernen am Modell... einen kontinuierlichen Prozeß dar, in dessen Verlauf neue Verhaltensweisen erworben und bestehende Verhaltensmuster bis zu einem gewissen Grade modifiziert werden, indem Einflüsse von verschiedenen aktuellen und symbolischen Modellen in allen Lebensperioden aufgenommen werden" (S. 68).

Unter Einbeziehung kognitiver Variablen wird eine Interdependenz von Verhalten, Person und Umwelt unterstellt:



So, wie die Umwelt auf das Verhalten von Menschen einwirkt, beeinflußt das Verhalten die Umwelt. Vermittelt werden diese Prozesse durch die Person als Interdependenz von Erwartung und Verhalten. Die Person steht nicht in der potentiellen, sondern in der durch das Verhalten beeinflußten tatsächlichen Welt.

Im Zentrum der Theorie von BANDURA steht die Frage nach der Entstehung und Aufrechterhaltung von Verhalten in dieser Interdependenz. Nicht primär die nachfolgenden Konsequenzen, sondern stellvertretende Erfahrung bestimmt die

psychologische Dimension des Lernens. Modellernen bildet den primären Mechanismus sowohl für die Aneignung (Akquisition) wie für die Ausführung (Performanz) von Verhaltenssequenzen. Dabei steuern vier Teilprozesse das Beobachtungslernen. 'Aufmerksamkeitsprozesse' beeinflussen sowohl die Hinwendung zur Beobachtung wie die exakte Wahrnehmung der wesentlichen Elemente der modellierten Handlung einschließlich der Merkmale des Modells. Diese wirken neben den Eigenarten der modellierten Tätigkeit, der Struktur der sozialen Interaktion und der Belohnungskontingenzen während der Beobachtung auf die Perzeptionsbereitschaft. Mit der 'symbolischen Repräsentation' führt BANDURA eine Anzahl kognitiver Variablen ein. Die dargebotene und beobachtete Verhaltenskette , einschließlich ihrer antezendenten und konsequenten Anreizstrukturen werden vom Beobachter ikonisch oder verbal kodiert und bei der späteren Performanz reaktiviert. Hierin unterscheidet sich das Beobachtungsverhalten kleiner Kindern von dem älterer Kinder bzw. Erwachsener. Die vom Kleinkind mögliche, unmittelbare Nachahmung stellt weniger Anforderungen an die kognitive Struktur als solche Nachahmungsreaktionen, die in Abwesenheit des Modells - zeitlich verzögert also - ausgeführt werden. Die perzipierte Information wird analog der Behaltensmodelle abgespeichert und kognitiv organisiert. Um sie vergessensresistenter zu machen, kann diese Information sowohl symbolisch wie motorisch geübt werden, d.h. sie wird entweder gedanklich oder durch Handeln wiederholt. Die 'motorische Reproduktion' setzt sowohl die physischen und psychischen Fähigkeiten voraus, um überhaupt reagieren zu können. Ebenso muß das neu zu erwerbende Verhalten an bereits beherrschte Teilfertigkeiten anknüpfen. Die Modifikation der eingeübten Verhaltenssequenzen bis zur Übereinstimmung mit dem Verhalten des Modells verläuft über Prozesse der Selbstbeobachtung und der Rückmeldung bis zur angemessenen Reproduktionsleistung. Wann das Verhalten realiter ausgeführt wird, hängt von der Motivstruktur des Beobachters ab. Die antizipierte Nützlichkeit, d.h. die vorweggenommene positive Wirkung auf die Umwelt und/oder auf die eigene Person bilden das auslösende Moment. Dabei wirken (wenn auch nach BANDURA nicht notwendigerweise; vgl BANDURA, 1977) drei mögliche Verstärkungsquellen: durch die Umwelt, durch die Person selbst (Selbstbekräftigung) sowie durch die gleichfalls abgespeicherte Kontingenz, der das Modell unterlag.

Die Modelldarbietung kann über drei unterschiedliche Modalitäten stattfinden. Entweder beobachtet der Rezipient ein physisches (in vivo) Modell. Dies gilt insbesondere für die vielfältigen 'unsystematisch' gelernten Verhaltensweisen durch Erwachsene und Peers. Zum anderen kann ein Verhalten verbal vermittelt werden. Hierunter fällen alle Ansätze zum Lernen durch Instruktion. Letzlich findet Lernen durch symbolische Modelldarbeitung statt. Diese modellierten Verhaltensweisen werden in der Regel durch Medien vermittelt. Modelle können dabei reale Personen wie auch Zeichentrickfiguren/-prozesse sein. Neben konkreten Verhaltensmustern sind auch abstrakte Modellierungen möglich. Dabei eignet sich der Beobachter grundlegende Prinzipien, d.h. dem Verhalten zugrundeliegende Regeln an. Mit Blick auf PIAGET und KOHLBERG verweist BANDURA (1977) auf die Genese des moralischen Urteils infolge abstrakter Modellierung.

Die dem Verhalten vorgelagerten und dem Verhalten folgenden Reaktionen bzw. Stimuli sind für die sozial-kognitive Lerntheorie, wenn auch in kognitiv-vermittelter Form von verhaltensbeeinflussender oder verhaltenserleichternder Bedeutung.

"Die Information über die wahrscheinliche Auswirkung bestimmter Handlungen oder Ereignisse werden durch Umweltstimuli übermittelt. Die Person wird darüber informiert, was sie von den besonderen Merkmalen der Orte, Personen oder Dinge bzw. der sozialen Zeichen in der Sprache, der Gestik und den Handlungen anderer zu erwarten hat" (BANDURA 1979, engl. 1977, S. 65).

Während die operante Konditionierung die Genese diskriminativer Stimuli durch das Paradigma der klassischen Konditionierung erklärt, betont die sozial-kognitive Lerntheorie die kognitive Vermittlung korrelierter Reizbeziehungen. Die Erwartung, die mit einem diskriminativen Stimulus verbunden ist, differenziert BANDURA (1977) in Handlungsergebnis- und Leistungseffizienzerwartung. Ein diskriminativer Stimulus evoziert bei einem Menschen die Kognition, daß ein bestimmtes Verhalten zu einem bestimmten, mit dem diskriminativen Stimulus angekündigten Ergebis führt (Ergebiserwartung). Erst die Überzeugung, dieses Verhalten auch mit hoher Wahrscheinlichkeit erfolgreich ausführen zu können (Leistungseffizienzerwartung) führt zur Ausführung des Verhaltens. Konsequenzen kontrollieren dabei zwar das Verhalten, dieses ist jedoch nicht von ihnen abhängig:

"So kann das Wissen um die Bekräftigungspläne einen größeren Einfluß auf das Verhalten ausüben, als die Bekräftigung selbst..." (BANDURA 1979, S.100f, engl. 1977).

Untersuchungen zeigen darüber hinaus, das Probanden sich weniger nach (konstanten) verabreichten Verstärkungsplänen, sondern nach verbal dargestellten (variablen) Verstärkungsplänen verhalten. Die Probanden reagierten auf das, was sie erwarteten und nicht nach den tatsächlichen Plänen (DULANY, 1968; KAUFFMAN, BARON & KOPP, 1966). Neben den externen Konsequenzen kontrollieren auch beobachtete wie interne Belohnungsstrukturen das Verhalten. Insbesondere der Einfluß der Selbstbekräftigung auf das Verhalten untermauert das Paradigma kognitiv kontrollierter Prozesse. Dabei unterscheidet BANDURA (1977) in Anlehnung an die Motivationsforschung drei Teilprozesse. Die Leistung definiert (1) die selbstgesetzten Bewertungsdimensionen, die (2) ihrerseits mit ebenfalls modellierbaren Standards bzw. Referenzleistungen verglichen werden können und (3) je nach Übereinstimmungsgrad positiv oder negativ selbst belohnt werden.

Daß die Belohnungsstruktur eine nicht unwesentliche Rolle spielt, zeigt FLANDERS (1968). Er differenziert zwischen vier verschiedenen Verstärkungsmöglichkeiten während des Modellernens: (a) weder das Modell noch der Beobachter werden verstärkt, (b) nur das Modell wird verstärkt, (c) nur der Beobachter wird (i) für seine Akquisition und/oder (ii) für seine Perfor-

manz verstärkt und (d) sowohl das Modell wie der Beobachter werden kontingent verstärkt. Unter Berücksichtigung unterschiedlicher Verstärkungsintensitäten während des Meßvorgangs gelangt er aufgund einer Zusammenfassung von Ergebnisbefunden zu folgender Schlußfolgerung:

1. Sowohl kontinuierliche wie partielle Verstärkung des Beobachters erhöhen das Nachahmungsverhalten des Beobachters
2. Wird statt des Beobachtungsverhaltens das Performanzverhalten während der Treatmentapplikation kontinuierlich oder partiell verstärkt, sinkt die Nachahmungsrate
3. Stellvertetende Verstärkung erhöht das Nachahmungsverhalten; je mehr Modellreaktionen verstärkt werden, desto mehr tendiert der Beobachter zur Nachahmung
4. Tentativ besteht ein Zusammenhang zwischen Verstärkungshöhe während der Messung der abhängigen Variablen und dem Umfang der Verstärkungskontingenzen während der Treatmentapplikation
5. die Experimente zeigen, daß Beobachter in Abwesenheit möglicher Verstärkungskontingenzen Modellverhalten nachahmen.

FLANDERS (1968) resümiert, daß die unter dem Begriff der 'behavioral contagiation' zusammengefaßten diskriminativen und/oder verstärkenden Mechanismen durch das Modell (per se) oder durch die Situation der weiteren Erforschung bedürfen. Mit BANDURA stimmt er darin überein, daß Verstärkungskontingenzen keine notwendigen, jedoch erleichternde Bedingungen darstellen (vgl. FLANDERS 1968).

Im Rahmen der Verhaltenstherapie wurden vielfältige Modifikationen auf der Basis des Modellernens durchgeführt. Zum (prophylaktischen) Modellernen pro-sozialer Verhaltensweisen liefert BRYAN (1975) einen Überblick. Von besonderem Interesse sind die Ergebnisse von BRYAN & WALBEK (1970), STAUB (1973), sowie YARROW et al. (1973). Soziale Modelle scheinen u.U. für prosoziale Verhaltensweisen weniger wirksam zu sein als zur Modellierung anti-sozialer Verhaltensweisen. In einem Experiment von ROSENHAN & WHITE (1967) führte erst die Einführung einer Wiederholungsphase zu einem 'langdauernden' pro-sozialen Verhalten.

2.2.2 Austauschtheoretische Erklärungen

Bereits SKINNER (1953) verweist auf Austauschaspekte in sozialen Interaktionen: Ein Mensch ist bereit zu geben, wenn er dafür etwas bekommt. Diese Leistung muß nicht auf gleicher Ebene wie die gebende Leistung liegen (vgl. STAUB 1979). Je weniger 'selbstlos' die gebende Leistung ist, desto externer werden die 'Gegenleistungen'. Sie variieren vom Spannungsabbau bzw. von der durch Selbstbekräftigung induzierten Freude über die Antizipation einer zukünftigen Entgegennahme notwendiger Hilfe, bishin zu öffentlichem Lob oder materieller Anerkennung. Über dieses universelles Grundparadigma informieren verschiedene theoretische Ansätze.

Die Austauschtheorie von THIBAUT & KELLEY (1953)

IRLE (1975) definiert eine soziale Interaktion als

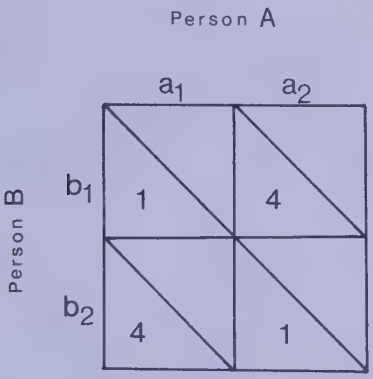
"... Aktionen einer Person P1, die Aktionen einer anderen Person P2 affizieren und umgekehrt P1 von Aktionen der anderen Person P2 affiziert wird" (S. 398).

Jede Handlung erfordert dabei Aufwand (Kosten) und liefert einen bestimmten Ertrag (Lohn). Die Theorie von THIBAUT & KELLEY (1953) ermöglicht die Analyse solcher sozialer Interaktionen unter dem Aspekt von Nutzen-Kosten-Beziehungen. Eine Interaktion als wiederholte, interpersonale Beziehung involviert Verhaltensweisen mit folgenden Merkmalen: (a) Präsenz der Teilnehmer, (b) gegenseitige Herstellung wechselseitiger Ergebnisse oder (c) wechselseitige Kommunikation (vgl. S. 10). Die kleinste Analyseeinheit besteht aus einer Verhaltenssequenz, die sowohl instrumentelles Verhalten - d.h. die zielgerichtete Bewegung einer Person - und rezeptives Verhalten - also wahrnehmendes und interpretierendes Verhalten - beinhaltet. Verhaltensweisen werden durch individuelle Anreize, Außeneinflüsse und durch die Verstärkungsstruktur für die Sequenz kontrolliert. Eine Interaktion führt zu zwei Konsequenzen, zum einen zur Belohnung in Form von Gratifikation, Vergnügen und/oder Triebreduktion und auf der anderen Seite zu Kosten, d.h. zu Faktoren, die das ausführende Verhalten hemmen: sei es in Form physischen und/oder mentalen Aufwandes, in Form von Konflikten und/oder kompetitiver Tendenzen. Die Nutzen oder die Belohnung kann als direkte Folge des eigenen Verhaltens und/oder des Verhaltens anderer verstanden werden. Die Autoren klassifizieren Kosten und Nutzen in endogene und exogene Determinanten. Endogene Faktoren ensprechen intrinsischen Motivationen, als exogene Variablen benennen THIBAUT & KELLEY (1953) soziometrische Wahlen, Fähigkeiten wie emotionale Stabilität, Pünktlichkeit, Humor, räumliche Nähe, Ähnlichkeiten der Einstellung, Komplementarität der Bedürfnisse sowie Belohnungsstärke und Status (vgl. S. 33 - 49). Verhaltensweisen können zudem mit der sozialen Interaktion kompatibel oder zu ihr interferent sein.

Die Interaktionspartner machen den weiteren Verlauf ihrer Interaktion vom Ergebnis (acceptance of outcome) in Relation zu alternativen Möglichkeiten (comparison level) abhängig. Eine Person ist in dem Maße von der sozialen Interaktion abhängig, wie das Ergebnis in positiver Richtung vom alternativen Vergleichs-Niveau abweicht. Verfügt ein Akteur über Fähigkeiten, auf die Qualität des Ergebnisses des Ko-Akteurs Einfluß auszuüben, verfügt dieser Partizipand über Macht (power). 'Macht zur Schicksalskontrolle' (fait control) liegt vor, wenn Person I in der Lage ist, das Ergebnis von Person II zu beeinflussen, unabhängig davon, was Person II tut. 'Macht zur Verhaltenskontrolle' (behavior control) liegt vor, wenn die Verhaltensänderung von I eine Änderung des Verhaltens von Person II notwendig macht. Wünscht beispielsweise ein Akteur, sein Ko-Akteur möge an einer kooperati-

ven Periode teilnehmen, und verfügt dieser Akteur über die Macht zur Verhaltenskontrolle, kann er sie zur Erreichung seines Ziels einsetzen:

Abb. 7: Verhaltenskontrolle als Möglichkeit zur Kooperation

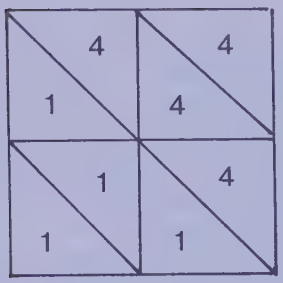


Akteur A möchte, daß sich sein Partner kooperativ (b(1)) verhält. Durch die Wahl a(2) kann er dieses Verhalten ausformen: immer, wenn Person B sich b(1) verhält, dann reagiert Person A mit dem Verhalten a(2) und immer, wenn sich Person B b(2) verhält, reagiert Person A mit a(1).

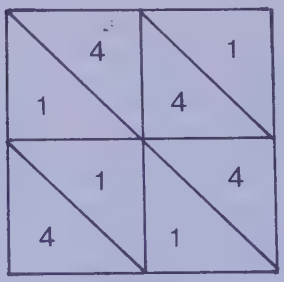
Interaktionen sind jedoch häufig interdependent. Dies zeichnet insbesondere (freiwillige) kooperative oder kompetitive Interaktionen aus.

Abb 8: Kooperative und kompetitve Interdependenz-Strukturen

Kooperative Interdependenz



Kompetitive Interdependenz



In einer kooperativen Interaktion erreichen beide Teilnehmer nur ein gutes Ergebnis, wenn sie die Kombinationen a(2)-b(1) zeigen. Unter kompetitiver Interdependenz können die Akteure nicht kooperieren: der Gewinn des einen ist immer mit dem (relativen) Verlust des anderen verbunden.

Soziale Interaktionen stehen zudem unter dem Einfluß von Regeln und der Aufgabenstruktur. Normen regulieren Dependenz zugunsten periodaler Interdependenz. So kann z.B. die kompetitive Interaktion durch Regel-Festlegung in eine kooperative Interaktion übergehen. Sprechen sich die Personen ab, wechselseitig b(2)-a(1) und b(1)-a(2) zu wählen, erlaubt dieser Reaktionsaustausch beiden Teilnehmern ein gutes Ergebnis. THIBAUT & KELLEY (1953) bezeichnen Normen und Regeln daher als ein drittes Agens:

"... the third agens exercises power over each member in the usual sense of making it desirable for him to act in certain ways at certain times and does so in an impartial way without a regard to the special interests of either one" (S. 129).

Die zu bewältigende Aufgabe innerhalb einer sozialen Interaktion betrachten die Autoren als einen Stimuluskomplex, durch den Personen offene und verdeckte Reaktionen äußern, die zu verschiedenen Ergebnissen führen. Aufgaben können nur eine (z.B. eine Mathematikaufgabe) oder mehrere Lösungen (z.B. die Lösung eines Problems) zulassen. Desweiteren unterscheiden THIBAUT & KELLEY (1953) zwischen konjunkten und disjunkten Aufgabenstellungen. Bei einer konjunkten Aufgabenstellung müssen beide Interaktionspartner Teilauf-

gaben erfüllen, um an einen Verstärker zu gelangen. Disjunkte Aufgaben sind unabhängig voneinander und ohne Koordination des gemeinsamen Ziels zu lösen.

Equity Theorie

Soll eine soziale Interaktion aufrecht erhalten bleiben, dann müssen Kosten und Nutzen gleich verteilt sein. Unter dieser 'Reziprozitätsnorm' (GOULDNER 1960) werden zwei Theoreme diskutiert. WALSTER und Mitarbeiter (WALSTER, BERSCHEID, & WALSTER 1970, 1973; WALSTER, WALSTER, & BERSCHEID 1978; WALSTER & WALSTER 1975; WALSTER & PELIAVIN 1972) liefern einen austauschtheoretischen Ansatz unter Einbeziehung dieses Reziprozitätsaspektes. Equity (HECKHAUSEN (1980) übersetzt dies mit 'Billigkeit') liegt vor, wenn Gruppenmitglieder ihr Ergebnis (output) im Verhältnis zu ihrem Beitrag (input) verteilen:

$$\frac{O(a) - I(a)}{I(a)^{**}k(a)} = \frac{O(b) - I(b)}{I(b)^{**}k(b)}$$

Die Exponenten k nehmen in Analogie zum Vorzeichen des Zählers die Werte +1 oder -1 ein. Das Problem im Umgang mit dieser Funktion liegt in der subjektiven Bewertung der einzelnen Terme.

Vor dem Hintergrund individuellen Strebens nach Gewinn können Gruppen dann einen gemeinsamen Gewinn maximieren, wenn die Kosten und die Nutzen zwischen den Gruppenmitgliedern 'equitabel' verteilt sind. Um dies zu ermöglichen, bilden Gruppen einen Sanktionsmechanismus: Mitglieder, die sich equitabel verhalten, werden belohnt, während die Mitglieder, die sich inequitabel verhalten, bestraft werden. Folgt einem inequitablen Verhalten keine negative Sanktion, sind Mitglieder geneigt, durch diese Strategie weiterhin ihr Ergebnis zu maximieren. Die Wiederherstellung von 'Equity' kann insgesamt auf zwei Ebenen erfolgen. Entweder verändert der Partizipand sein Verhalten (durch Reduktion des Inputs oder durch den Versuch, seinen Gewinnanteil zu erhöhen), oder der 'ausgebeutete' Partizipand versucht, psychologische Equity herzustellen, indem er seine Wahrnehmung bezüglich seines oder bezüglich den Input und/oder Output des Ko-Akteurs adjustiert.

Dem Equity-Konzept steht das Equality-Konzept gegenüber. SEILER (1978) unterscheidet soziale Interaktionen nach Aufgabenorientierung und interpersonaler Orientierung. Aufgabenorientierte Gruppen orientieren sich am Equity-Prizip, während solche Interaktionen, die auf Solidaritätsaspekten beruhen, sich nach dem Equality-Prizip orientieren. Empirische Evidenz hierfür liefern nach SEILER (1978) vor allem DEUTSCH (1975), LEVENTHAL (1976a, 1976b) und SAMPSON (1975). LERNER (1977) bestätigt diese Unterscheidung und formuliert darüberhinaus die Hypothese, daß Verstärkung in Abhängigkeit der Leistung das Equity-Prinzip aktiviert und Verstärkung nach

dem Grundsatz 'lump sum' zur Bevorzugung des Equality-Prinzips führt. Bei identischem Input der Ko-Akteure gehen die beiden Prinzipien ineinander über.

2.2.3 Das Modell von HAKE

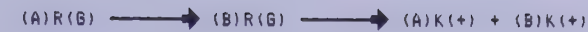
Einen Systematisierungsversuch, mit dessen Hilfe die prozeduralen Dimensionen, die Kooperation beeinflussen, integriert werden sollen, elaborieren HAKE & VUKELICH (1972). Dabei steht die erklärungsleitende Frage, "What can we do to produce cooperation?" (HAKE & OLVERA, 1978, S. 20) im Mittelpunkt. Eine Kooperation wird nach diesem Modell durch die Reziprozität von Reaktions-Verstärkungs-Abhängigkeiten ausgezeichnet. Die Interktionspartner ermöglichen durch ihre gebenden Reaktionen primär die Verstärkung des Ko-Akteurs, der seinerseits durch seine gebenden Reaktionen den Akteur verstärkt:

```
Person A:  R(G) ──→ (B) R(P) ──→ (B) K(+)   1. KE
-------------------------------------------------- KP
Person B:  R(G) ──→ (A) R(P) ──→ (A) K(+)   2. KE
```

R(G) = gebende Reaktion, (B) R(P) instrumentelle Reaktion der Person A, (B) k(+) kontingente Verstärkung für Person B, 1. KE = erste kooperative Episode, KP = kooperative Periode.

Die Grundstruktur einer kooperativen Periode läßt bereits die Kontingenzkombinationen erkennen. Eine interdependente Kontingenz besteht, wenn der Verstärker einer Person sowohl vom Verhalten der anderen Person wie vom eigenen Verhalten abhängig ist. In der Definition zur Kooperation (vgl. Abschnitt 2.1) wurde die 'zumindest teilweise Dependenz' in der Verstärkungskontingenz konstatiert. Komplette Dependenz besteht, wenn der Verstärker eines Teilnehmers vollkommen unter der Kontrolle des Ko-Akteurs steht (vgl. auch BOREN, 1966). Eine asymmetrische Dependenz liegt vor, wenn die Verstärker der Teilnehmer vollkommen von nur einem Mitglied der sozialen Interaktion abhängt. Partielle Asymmetrie liegt entsprechend vor, wenn der Verstärker des Partners komplett unter der Kontrolle des Ko-Akteurs und der Verstärker des Ko-Akteurs von beiden Teilnehmern abhängt. Alternierende Asymmetrie bezeichnen HAKE & OLVERA (1978) als 'mixed contingency' (S. 223).

Reziprozität muß nicht unbedingt in jeder Periode realisiert werden. Ist sie über mehrere Perioden möglich, wird dies in Analogie zu THIBAUT & KELLEY als Reaktionsaustausch bezeichnet. Als Reaktions-Sharing wird eine interdependente Kontingenz bezeichnet, durch die beide Teilnehmer innerhalb einer Episode eine Belohnung erhalten:



Eine Vielzahl von Kooperationsexperimenten bedienen sich insbesondere dieser Kontingenzstruktur (vgl. weiter hinten).

Die Kosten variieren während einer kooperativen Sequenz infolge der näheren Rahmenbedingungnen. Liegen unbegrenzte Verstärkungs-Ressourcen vor, kostet den Akteur eine gebende Reaktion solange nichts, wie der Ko-Akteur zur Kooperation bereit ist. Bei begrenztem Ressourcenumfang oder bei ergänzenden Bestrafungskontingenzen können hohe Kosten entstehen (vgl. weiter unten die Darstellungen zum PDG). Eine weitere Dimension liegt in der Verstärkungskontingenz für alternative Reaktionen. Häufig werden alternative Reaktionen in Experimenten, die Handlungsinstrumente verwenden, nicht belohnt. Daher fragt BRYAN (1975) nicht ohne Ironie, ob in solchen Experimenten Kooperation nicht als Artefakt betrachtet werden sollte.

2.2.4 Das Kooperationskonzept von M. DEUTSCH

DEUTSCH (1949, 1962, 1965) entwickelt eine Kooperationstheorie im Rahmen der Feldtheorie von LEWIN (1939). Das Interdependenzmodell der Feldtheorie erlaubt dabei keine 'stringente' Unterscheidung zwischen abhängigen und unabhängigen Variablen. Auf ein Ziel gerichtete Verhaltensweisen werden vielmehr durch einen multivariaten Pfad beschrieben. Aus der Gestaltpsychologie übernimmt LEWIN die Beschreibung und holistisch-strukturelle Erklärung von Phänomenen (vgl.SOLLE 1969). Sie können aus dem Feld heraus erklärt werden.

"...field, that is, as part of the totality of coexisting facts which are conceived as mutually interdependent" (DEUTSCH 1968, S.414).

Psychologische Prozesse sind dabei nicht phänotypisch-klassifizierend, sondern genotypisch-konstruierend zu erfassen. Phänotypische Daten mussen in die Sprache ihrer Konstrukte übersetzt werden, bevor es gelingt, psychologische Gesetzmäßigkeiten, die auf den Einzelfall anwendar sind, zu entwickeln. Mit dem 'Lebensraum' versucht LEWIN der Realität soweit als möglich nahezukommen, indem er die Interdependenz von Person und Situation aufgreift .

"LEWIN'S emphasis on the relation of the concrete individual to the concrete situation leads the scientific investigator to a more explicit realization that the understanding of behavior requires not only a knowledge of his immediate situation" (DEUTSCH 1968, S. 418).

Personen tendieren dazu, ein dynamisches Equilibrium zu ihrer Umwelt aufrechtzuerhalten. Der Lebensraum umfaßt dabei einen Komplex der das Verhal-

ten determinierenden Faktoren. Es setzt sich aus der Interaktion der Person mit der von ihr wahrgenommenen Umwelt zusammen:

Verhalten = f (Lebensraumes) = f (Person, Umwelt)

Der 'Umwelt'-Begriff wird bei LEWIN ambivalent benutzt: zum einen bezeichnet er damit die objektive Umwelt in Form von Stimulus-Situationen, ein anderes Mal ist damit die psychologische Umwelt, so wie sie für das Individuum existiert, gemeint (vgl. DEUTSCH, 1968, S. 423). Nicht weniger eindeutig ist der Begriff der 'Person'. Einmal bezeichnet LEWIN damit die mit der Person verbundenen Merkmale, ein weiteres Mal die Relation der Person in ihrem Lebensraum, und letztlich als das sich verhaltende Selbst, d.h. als die vom Individuum wahrgenommenen Beziehungen zwischen sich und seinem Lebensraum.

Das topologische Konzept des Raumes beinhaltet drei Bedingungen, die Region, die Position und die Lokomotion. Eine Region umfaßt die aktuellen Aktivitäten und nicht die objektiven Flächen, in denen diese Aktivitäten stattfinden. Die Position der Person wird durch die Aktivitätsregion, in der die Person steht, definiert. Lokomotion bezeichnet dabei die Bewegung der gegenwärtigen Position zu einer Nachbarregion. Aneinandergereihte Regionen bilden einen Pfad. Die Bewegungsrichtung der Person wird durch das dynamische Konzept, zu dem die Begriffe Spannungssystem, Kraft und Valenz gehören, erklärt. Spannung liegt vor, wenn im Individuum ein Bedürfnis (need) oder eine Intention (quasi-need) vorherrscht. Die Spannung wird durch Befriedigung des Bedürfnisses oder der Intention abgebaut. Die Kraft charakterisiert dabei die Richtung und die Stärke der Veränderungstendenz. Analog zum Kräfteparallelogramm in der Physik bilden kombinierte Kräfte eine Kräfte-Resultante. Wenn diese existiert, findet Lokomotion entweder im Verhalten oder in der kognitiven Struktur statt. Für jede Region gibt es einen bestimmten Wert, diese auch zu erreichen. LEWIN verwendet hierfür den Begriff der Valenz.

Kooperation beinhaltet die Zielinterdependenz von Personen. DEUTSCH (1949) leitet daraus verschiedene Implikationen ab:

1. Es gibt Ziele, die allein durch eine gemeinsame Zielerreichung realisierbar sind (promotively interdependent goals).

   "If A,B,C, etc, does not obtain his goal (enter his goal region), X does not obtain his goal. X obtains his goal only if A,B,C, etc., obtain theirs. A, B, C, etc. obtain their goals only if X obtains his" (DEUTSCH 1949, S. 133).

2. Ziele sind konträr zueinander, wenn ein Teil der Personen (bzw. eine Person) nur ihr Ziel erreichen kann, während ein anderer Teil von Personen (bzw. eine andere Person) ihr Ziel nicht erreicht und vice versa.

3. Eine Zielstruktur ist individuell, wenn keine Interdependenz für die Zielerreichung besteht.

Die unterschiedlichen Relationen können als Wahrscheinlichkeitsmatrizen dargestellt werden (DEUTSCH 1962, S. 277):

Abb. 9: Wahrscheinlichkeiten der Zielerreichung von zwei Personen (nach DEUTSCH 1962, S. 277)

FÖRDERNDE

Person 1 \ Person 2	G	$\bar{G}$
G	p	0
$\bar{G}$	0	1 − p

KONTRÄRE

Person 1 \ Person 2	G	$\bar{G}$
G	0	1 − p
$\bar{G}$	p	0

FEHLENDE

Person 1 \ Person 2	G	$\bar{G}$
G	$p_1\ p_2$	$p_1[1-p_2]$
$\bar{G}$	$p_2[1-p_1]$	$[1-p_1][1-p_2]$

INTERDEPENDENZ

Die Wahrscheinlichkeit ist sowohl eine Funktion der Verhaltenswahlen (B), der Akteure wie der Ereignisse, die außerhalb der Kontrolle der Akteure liegen. Ein Individuum muß zunächst überhaupt wahrnehmen, welche Form der Interdependenz besteht. Dies bedeutet, die Teilnehmer an einer (interdependenten) Interaktion müssen zunächst die Existenz des anderen perzipieren sowie die Leistungen des anderen als Beitrag zur Erreichung des eigenen Ziels interpretieren können.

Die Bereitschaft, den anderen zu fördern, wird durch fünf Aspekte beeinflußt: (1) durch die Wichtigkeit des Ziels, (2) durch den wahrgenommenen Ertrag, der aus solchem Verhalten bzw. der Zielerreichung erwachst, (3)

durch die wahrgenommene Wahrscheinlichkeit, das Ziel auch erreichen zu können, (4) durch die wahrgenommene intrinsische Attraktivität und der fehlenden Kosten für die notwendigen Aktionen, sowie (5) durch die wahrnehmbare Unmittelbarkeit der Zielerreichung infolge der Aktion (DEUTSCH 1962 S. 278). Diese Bedingungen beeinflussen zunächst nur die Entscheidung, den kooperativen Pfad zu gehen. Die hiermit verbundene Ambiguität dieses Pfades erlebt der Akteur als Vertrauensproblem. Ist die negative Valenz größer als die positive, muß eine vertrauensvolle Wahl getroffen werden. Vertrauenswahl aus Zuversicht liegt vor, wenn der Akteur annimmt, daß der erwünschte und nicht der unerwünschte Effekt eintritt. Desweiteren vertraut die Person auf das Eintreten von Reziprozität. Reziprozität ist in solchen Situationen, in denen der Akteur sein Ziel nicht ohne den Ko-Akteur erreichen kann, wichtiger als in Situationen, die Kooperation erlauben, jedoch nicht erfordern.

In nicht-erforderlichen Kooperatonssituationen beeinflussen weitere Variablen die Entscheidung zur Kooperation:

(1) die Attraktivität der mit der Kooperation verbundenen Erfahrung
(2) die Attraktivität des Ko-Akteurs
(3) die Attraktivität, als Gruppenmitglied identifiziert zu werden
(4) die Attraktivität, durch die Kooperation das Ziel leichter erreichen zu können und
(5) die Unmittelbarkeit, mit der der kooperative Pfad zum Ziel führt (DEUTSCH, 1962, S. 294f).

Mit Hilfe dieser Determinanten sowie der Wahrnehmung der Fertigkeiten und der Ressourcen des Ko-Aktuers bzw. der Wahrnehmung der subjektiven Wahrscheinlichkeit, daß diese Merkmale auch eingebracht werden, bildet der Akteur seine subjektive Wahrscheinlichkeit, mit der das kooperative Engagement zum Erfolg führen wird.

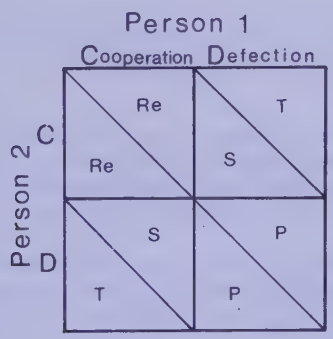
Die Bedingungen des Vertrauens und der Attraktivität sind um die Probleme der Verhandlung (bargaining) und die der Koordination erweitert. Das Verhandlungsproblem ergibt sich aus der Reaktions-Austausch-Kontingenz: beide Teilnehmer erreichen ihr Ziel nur dann, wenn sie eine Übereinstimmung ihrer Schrittfolgen erreichen. Diese Übereinstimmung wird durch die Interessen der Teilnehmer bestimmt, kooperieren oder kompetitieren zu wollen. Dies beinhaltet, über die Ressourcen und die Freiheit der Teilnehmer zu verhandeln. Koordination der Aktionen ist insbesondere dann notwendig, wenn mehrere Verhaltensalternativen vorliegen.

2.2.5 Das Gefangenendilemma als Forschungsparadigma

Kooperationsverhalten wird häufig an Untersuchungsinstrumenten studiert, die auf Nicht-Nullsummen-Spielen basieren. Nicht-Nullsummen- und Nullsummen-Spiele gehören zur Kategorie der Wahlinstrumente und finden im Rahmen der Konfliktforschung auf der Grundlage von Spieltheorien Anwendung (vgl. RAPOPORT & OWANT 1962, S.1). Die Teilnehmer - in der einfachsten wie häufigsten Anwendung Dyaden mit jeweils zwei Wahlmöglichkeiten - können gegeneinander oder miteinander spielen. Bei Nullsummenspielen gewinnt der eine, was der andere verliert: die Summe von Gewinn und Verlust ist Null. Bei Nicht-Nullsummen-Spielen gewinnt der eine mehr als der andere, der somit relativ verliert.

Das Gefangenen-Dilemma-Spiel (PDG) ist das bekannteste wie in der Kooperationsforschung häufigst angewandte Experimental- und Meßinstrument. Sein Name geht nach RAPOPORT (1964) auf eine Erzählung von TUCKER zurück:

Zwei Gefangene sitzen unter dem Verdacht ein, eine Tat gemeinsam ausgeführt zu haben. Dem Staatsanwalt fehlt es an Beweisen, so daß er jedem einzelnen verspricht, jenen freizulassen, der als erster die Tat eingesteht. Die getrennt inhaftierten Gefangenen stehen vor einem Dilemma: schweigen sie beide, müssen sie entlassen werden; befürchtet jedoch einer von beiden, der jeweils andere Verdächtigte könnte ein Schuldbekenntnis ablegen, so gilt es, vor ihm dies zu tun. Jedoch ist das Bekenntnis nur sinnvoll, wenn der andere schweigt. Dieses Dilemma wird mit Hilfe einer Wahl- und Auszahlungsmatrix dargestellt.



Die Zellenindices bedeuten in der Terminologie RAPOPORT & CHAMMAH (1965):

Re = REWARD: die Belohnung für das gemeinsam eingebrachte Vertrauen
S = BUCKER'S PAYOFF : Verlust des betrogenen Ko-Akteurs
T = TEMPTATION: die Versuchung zu mogeln
P = PUNISHMENT: die Strafe für das gemeinsam entgegengebrachte Mißtrauen

Wählen beide Spieler (C), werden sie für ihre Kooperation belohnt, bzw. sie erhalten den höchsten gemeinsamen Gewinn, denn es gilt für PDG-Matrizen: Re gt S, (T+S)/2, P. Spielt der eine Akteur (D) und der andere Akteur (C), wird der eine für seine nicht-kooperative Wahl belohnt (T gt Re) und der andere für seine Kooperationsbereitschaft bestraft (P le Re, T). Die rationale Strategie liegt jedoch nicht im Reaktionsaustausch, da (T + S) / 2 le Re beträgt, sie liegt auf lange Sicht in der Kooperationsstrategie, es sei denn, ein Teilnahmer ist bereit, sich systematisch ausbeuten zu lassen.

Die Variation der Zellenwerte und damit des Konflikts führte in der Entwicklung der Konfliktforschung zu einer Vielzahl von PDG-Experimenten bzw. PDG-Varianten (vgl. RAPOPORT & CHAMMAH 1965, RAPOPORT & ORWANT 1962, KRIVOHLAVY 1974). Die Heterogenität der Ergebnisse und das niedrige Kooperationsniveau, motivierten zu dem mathematischen Spiel, durch vielfältige Indices die Ergebnisvariation zu 'erklären' (vgl. RAPOPORT & BUYER 1966).

VINACKE (1969) versucht, die vielfältigen Variablen in ein theoretisches Konzept zu integrieren. Er unterscheidet zwischen Einflußquellen durch die Aufgaben, durch die Situation und durch die Akteure. Die Aufgabe, und dies meint im PDG die Werte-Verteilung in den Zellen der Spielmatrix, beeinflussen das Spiel ebenso wie die Modalitäten der Präsentation des Spiels. Die Länge bzw. die Anzahl der Spielzüge führt nach seiner Übersicht zu einer Veränderung der Wahlentscheidungen. Situationbedingungen können besonders gut durch die festgelegten Spielstrategien des Simultanspielers variiert werden.

Kommunikation und die Instruktion beeinflussen die Wahl entscheidend. Als Persönlichkeitsmerkmale wurden neben der Geschlechtsvariable und dem Alter der Spieler häufig auch Einstellungen gemessen. Sie beeinflussen ebefalls die Intention der Spieler.

VINACKE (1969) versucht, die Vielfalt der Ergebnisse mit Hilfe der Feldtheorie zu systematisieren, da nach seiner Ansicht Menschen weder im ganzen rational agieren (wie die Spieltheorie dies unterstellt) noch durch antezedente intrinsische Merkmale allein geleitet sind:

"a need to look comprehensively at the interaction between person and environment, identifying variables from both directions, and seeking to determine how they are related in producing behavior" (S. 314).

Eine exemplarischen Untersuchung:
Die Studie von GUMPERT, DEUTSCH & EDWARDSON (1969)

Anliegen der Untersuchung:
In einem PDG-Spiel mit mehreren Durchläufen (Zügen) soll die Hypothese getestet werden: Je höher der reale Bezug zum Gewinn bzw. Verlust ist, desto mehr kooperative Wahlen finden statt.

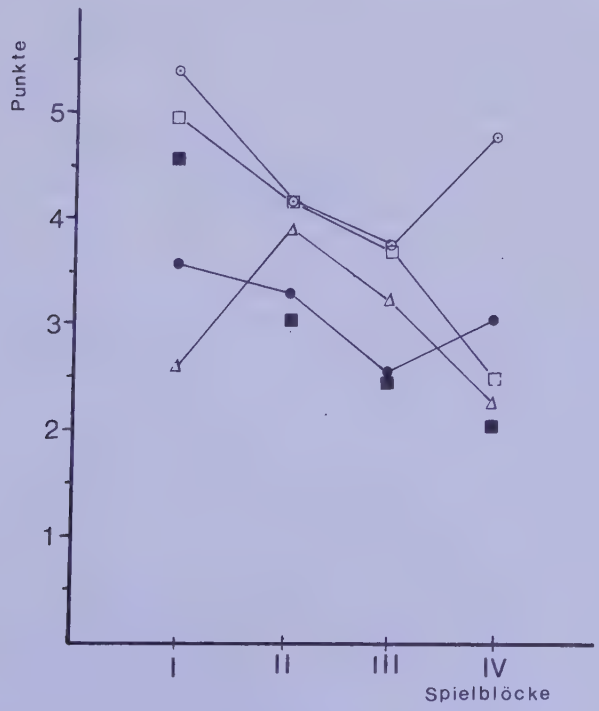
Verfahren:
Der reale Bezug (incentive magnitude) wurde als unbhängige Variable in fünf Stufen variiert. Je eine Gruppe spielte mit

(a) imaginären Dollars (ID) ⊙
(b) realen Dollars (RD) ■
(c) dem Eintauschwert 1 Punkt = 1 Cent (1) □
(d) dem Eintauschwert 1 Punkt = 5 Cent (5) △
(e) dem Eintauschwert 1 Punkt = 10 Cent (10) ●

Für die Marix wurden die folgende Werte eingesetzt:
Re = +1 T = +2 S = -2 P = -1

Ergebnisse:
Die Varianzanalyse weist keinen Effekt zwischen den fünf Treatmentvariationen aus. Aufgezeigt werden kann demgegenüber ein Effekt über die Anzahl der Züge (Durchlaufeffekt). Der Vergleich zwischen ID und RD weist auf einen signifikanten Treatmenteffekt in einer der Erwartung entgegengesetzten Richtung.

Abb. 10: Graphische Darstellung der Ergebnisse



Interpretation

Die Ergebnisse zeigen, daß die Höhe des Realitätsbezuges die Kooperation nicht positiv beeinflußt. Die Ergebnisse zeigen weiter auf, daß mit zunehmenden Spielzügen die Kooperation sinkt.

TEIL II

DURCHFÜHRUNG DER META-ANALYSE

3. PROBLEMSTELLUNG

Die vorliegende Meta-Analyse integriert erstmals eine Vielzahl empirischer Untersuchungen, in denen kooperatives Verhalten auszuformen versucht wird. Die Untersuchungen beziehen sich im wesentlichen auf vier Theorien: der Theorie des operanten Konditionierens, der Theorie des Modellernens, der Austauschtheorie und der Feldtheorie nach DEUTSCH. Dies bringt vielfältige Unterschiede zwischen den Studien nicht nur auf der methodologischen Ebene oder vor dem Hintergrund unterschiedlichster Stichproben (Populationen) mit sich, sondern auch und insbesondere auf der Ebene des Treatments. Ziel der Meta-Analyse ist es,

1. die Effektgrößen in Abhängigkeit der Referenztheorien und der mit ihnen verbundenen Konstrukte zu ermitteln,

2. die Effektgrößen auf der Ebene funktional-analytisch transformierter Variablen zu vergleichen,

3. den Einfluß methodisch und methodologisch bedingter Unterschiede (z.B. Dauer des Treatments, Art des Meßinstruments, Studienvalidität) zu kennzeichnen,

4. mit Hilfe regressionsanalytischer Verfahren heuristische Hinweise für wesentliche Einflußfaktoren auf Kooperation im Rahmen der Untersuchungen zu liefern.

Unterschiedliche Theorien involvieren unterschiedliche Hypothesen und damit unterschiedliche Treatmentapplikationen, die auf den ersten Blick inkommensurabel erscheinen. Demnach würde eine Meta-Analyse auf dem Vergleich formal beschreibarer Unterschiede stehen bleiben, ungeachtet der Frage, ob diese Unterschiede auch auf der operationalen Handlungsebene vorzufinden sind. Diese Unterscheidbarkeit ist dann kein Problem, wenn solche Unterschiede sich physikalisch manifestieren, wie z.B. in der Meta-Analyse von LYNN & DONOVAN (1980) in der sie die Wirkung medikamentoser gegenüber operativer Herztherapie untersuchen. Je weniger eindeutig jedoch physikalische Unterschiede sind, desto weniger eindeutig sind die prädiktiven Variablen auf der operationalen Handlungsebene zu differenzieren. COOK & CAMPBELL (1979) bemerken hierzu, daß den Daten ihre Erklärung - einmal den Einfluß von Hintergrundvariablen unbeachtet gelassen - gleichgültig ist. Will man dies übertragen, dann ist es dem empirisch ermittelten Kooperationswert egal, ob er durch unmittelbare Bekräftigungskontingenzen oder durch Vertrauen erzeugt wurde.

Mit Hilfe eines Kodiersystems für alle ausgewählten Studien wird daher nachfolgend das methodische Vorgehen in der Treatmentvergabe bezüglich Bekräftigungskontingenzen und diskriminativer Stimuli umformuliert. Das Paradigma der operanten Konditionierung wird aus zwei Gründen gewählt: zum

einen erlaubt es durch seine Sparsamkeit, auf Konstrukte oder intervenierender Variablen zu verzichten. Auf der anderen Seite beinhaltet es zwei Dimensionen, die in der Diskussion um die Ausformung kooperativer Verhaltensweisen unabhängig der theoretischen Provenienz wiederholt angeführt werden: den Einfluß von Bekräftigungskontingenzen in der interdependenten Interaktion sowie diskriminative Hinweise als 'Voreinstellung zur Kooperation. Neben den transformierten Variablen 'Kontingente Bekraftigung kooperativer Reaktionen VOR dem Meßvorgang', 'Kontingente Bekräftigung alternativer Reaktionen VOR dem Meßvorgang', 'Diskriminative Hinweise VOR dem Meßvorgang', 'Kontingente Bekräftigung WÄHREND des Meßvorgangs' und 'Diskriminative Hinweise WÄHREND des Meßvrogangs' wurden weitere Bedingungen kodiert, durch die sich spezifische Diskriminationsmöglichkeiten oder Bekräftigungsstrukturen manifestieren. Die Variable 'Modellernen' beinhaltet abgestuft verschiedene modellierte Kooperationstrategien und Kooperationskontingenzen. 'Tit-for-Tat' kodiert die Applikation einer in der Literatur umfassend diskutierten kontingenten Kooperationsstrategie. 'Display informiert darüber, ob während des Meßvorgangs Informationen zum Spielerfolg den Interaktionspartnern mitgeteilt werden. Die Variable PWK kodiert den Prozentwert kooperativer und unkontingenter Wahlen durch den simultanen Ko-Akteur, während 'Revenuen' das Niveau der Eintauschverstärker differenziert. Es wird angenommen, daß mit Hilfe der umformulierten Variablen Unterschiede zwischen den Effektgrößen erkannt und beschrieben werden können.

In Abhängigkeit unterschiedlicher Untersuchungsmerkmale werden dabei Erklarungsmodelle generiert. Sie bilden die Grundlage für die Interpretation der Ergebnisse vor dem Hintergrund der in und zwischen den Untersuchungen diskutierten Probleme bzw. Hypothesen. Die Effektgröße interessiert dabei weniger in ihrer 'absoluten' Ausprägung.

4. Vorgehensweise

Die statistische Datenintegration umfaßt eine Folge systematischer Schritte. Zunächst wird der Gegenstandsbereich mit Hilfe von Handbuchern, Klassikern' und Übersichtsartikeln umrissen. Im zweiten Schritt sind die potentiell zu integrierenden Studien zu suchen. Mit Hilfe des Thesaurus des ERIC und der Psychological Abstracts können die wesentlichen Suchbegriffe entwickelt werden. Im vorliegenden Fall wurden für Kooperation als abhangige Variable über verschiedene Suchschritte (Psychological Abstracts, ERIC, , Dissertation Abstracts und vorliegende Untersuchungen) n = 220 Einzelstudien gefunden. Folgende deutschsprachigen Zeitschriften wurden separat durchgesehen:

Tabelle 13: Deutschsprachige Zeitschriften zur Kooperationsforschung

Psychologische Beiträge	1965 - 1981
Psychologische Praxis	1969 - 1981
Psychologische Rundschau	1970 - 1981
Zeitschrift für angewandte Psychologie	1960 - 1981
Zeitschrift für Differentielle und Diagnostische Psychologie	1980 - 1981
Zeitschrift für empirische Pädagogik	1977 - 1982
Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie	1969 - 1981
Zeitschrift für erziehungswissenschaftliche Pädagogik	1967 - 1981
Zeitschrift für klinische Psychologie und Psychotherapie	1972 - 1976
Zeitschrift für Psychologie	1969 - 1982
Zeitschrift für Sozialpsychologie	1970 - 1981

Diesen Zeitschriften konnten 12 Beiträge unter der formulierten Fragestellung entnommen werden. In einer ersten Vorprüfung wurden diese Arbeiten gesichtet und nach folgenden Kriterien selektiert, die auf etablierten Gültigkeitsüberlegungen basieren:

(1) in den Untersuchungen werden Variablen MANIPULIERT: Untersuchungen, in denen 'presage'- oder 'context'-Variablen untersucht werden, fielen heraus

(2) die Berechnung einer Effektgröße mußte möglich sein: es entfielen solche Beiträge, in denen z.B. der mitgeteilte

F-Wert auf multiple Faktoren zurückgeführt werden kann
und weitere Statistiken nicht vorlagen

Anhand dieser Kriterien reduzierte sich die Menge der Einzelstudien auf n = 49 Beiträge, die den folgenden Quellen entnommen wurden:

Tab. 14: Verteilung der Studien nach Publikationsherkunft

Art und Quelle der Veröffentlichung	Anzahl der Studien	cum
Zeitschriften	(36)	
Journal of Personal Social Psychology	7	7
Journal of Conflict Resulution	6	13
Journal of Abnormal and Social Psychology	3	16
Psychological Reports	3	19
Child Development	2	21
Journal of Advanced Behavior Analysis	2	23
Journal of Expected Child Psychology	2	25
Journal of Psychology	2	27
Zeitschrift für Sozialpsychologie	2	29
AVCR	1	30
Developmental Psychology	1	31
Elementary School Journal	1	32
Journal of Expected Social Psychology	1	33
Journal of Genetic Psychology	1	34
Journal of Personality	1	35
Journal of Social Psychology	1	36
Psychonomic Science	1	37
Dissertationen	11	48
Buch-Veröffentlichung	1	49

Nach wiederholter Durchsicht dieser Beiträge wurde ein Kodiersystem entwickelt. Aus den in Tabelle 8 dargestellten Veröffentlichungen konnten n = 156 Kodieritems extrahiert werden. Anhand aktueller Literatur zu Design- und Validitätsaspekten (z.B. COOK & CAMPBELL, 1979) wurden weitere Kodieritems entnommen. Letztlich mußten für den spezifischen Gegenstandsbereich 'Kooperation' die entsprechenden Items konstruiert werden. Für das Kodiersystem ergaben sich die folgenden Kategoriengruppen:

Identifikation der Untersuchung und des Effekts
Publikationsmerkmale
Theoriegeleitete Items
Information über die unabhängigen Variablen
Information über die abhängige Variable

Methodische Qualität der Studie
Statistische Basis der Effektberechnung
Berechnung der Effektgröße.

Die hierunter subsummierten Items wurden anhand von vier Studien auf ihre Anwendbarkeit hin überprüft. Tabelle 15 zeigt in der Übersicht die für die Meta-Analyse verwandten Items.

Tabelle 15: Die Items des Check-Systems für die Meta-Analyse (die mit * gekennzeichneten Items werden in der Folge umfassend beschrieben)

Iteminhalt	Anzahl der Kodierungen
IDENTIFIKATION DER UNTERSUCHUNG UND DES EFFEKTS	
Identifikationcode der Effektgröße	259
Laufende Nummer der Effektgröße in der Meta-Analyse	259
Laufende Nummer der Untersuchung in der Meta-Analyse	49
Datum der Kodierung	259
PUBLIKATIONSMERKMALE	
Publikationsjahr	259
Publikationsart (Zeitschriften, Dissertation, Monographie)	259
Publikationsumfang (in Seiten)	259
Einzelfall- vs. Gruppenstudie	259
Anzahl der Effektgrößen pro Studie	259
THEORIEGELEITETE ITEMS	
Explikation der Fragestellung (ja/nein)	259
Kategorie der Fragestellung	* 234
Explikation des Problems (ja/nein)	259
Explikation der Problemstellung	* 259
Referenztheorie	* 256
Referezkonzept	* 240
Literaturumfang	259
Anzahl inhaltlich formulierter Hypothesen	259
Anzahl statistisch formulierter Hypothesen	259
INFORMATION ÜBER DIE UNABHÄNGIGEN VARIABLEN	
Explikation der UV (ja/nein)	259
Operationalisierung der UV	259
Modellernen (UV-Set MOD)	* 42
Bekräftigungs-Kontingenzen (UV-Set KBV)	* 34

Fortsetzung von Tab. 15

Iteminhalt	Anzahl der kodierungen
Bekräftigungs-Kontingenen alternativer Reaktionen (UV -Set KBI)	* 22
Diskriminative Hinweise (DHV)	* 89
Reaktionskontingenzen UV-Set KBW)	* 200
Tit-for-Tat (UV TFT)	* 23
Prozentsatz kooperativer Wahlen des simultanen Mitspieler	* 39
Display des Kooperationsverhaltens	* 22
Diskriminative Hinweise während des Meßvorgangs (UV-Set DHW)	* 185
Revenuen der Verstärker	* 195
Zellenwert Re	* 182
Zellenwert T	* 182
Zellenwert S	* 182
Zellenwert P	* 182
Instruktion	* 97
Treatmentdauer in Minuten	76
Treatmentdauer als Anzahl der Spielzüge	154
Konfundierung des Treatments und Meßvorgangs	259
DIE KODIERUNG DER ABHÄNGIGEN VARIABLE	
Definition der AV (ja/nein)	70
Klassifikation der AV	* 252
Meßaggregat der AV	* 236
Reliablität der AV	69
POPULATION UND STICHPROBE	
Nennung der Population	23
Stichprobengröße	259
Umfang der Treatmentgruppe	233
Umfang der Kontrollgruppe	232
Anzahl genannter Stichprobenmerkmale	259
Grad der Kompetition	253
Alter der Probanden	245
Geschlecht der Probanden	245
Sozioökonomischer Status der Probanden	65
Stadt- vs. Landbevölkerung	92
METHODISCHE EVIDENZ DER STUDIE	
Stichprobenverfahren	240
Zuordnung der Gruppen	168
Validität des statistischen Schlusses	* 259
Konstruktvalidität	* 259
Interne Validität	* 259
Externe Validität	* 259
Übereinstimmung zwischen Daten und Diskussion	259

Fortsetzung von Tab. 15

Iteminhalt	Anzahl der Kodierungen
Transparenz der Untersuchung	259
STATSTISCHE BASIS DER EFFEKTBERECHNUNG	
Datenbasis des Effektwertes	* 259
Kalkulation der Effektgröße	* 259
Quelle des arithmetischen Mittels	* 259
Arithmetisches Mittel der Treatmentgruppe im Vortest	28
Standardabweichung der Treatmentgruppe im Vortest	21
Arithmetisches Mittel der Kontrollgruppe im Vortest	5
Standardabweichung der Kontrollgruppe im Vortest	5
Arithmetisches Mittel der Treatmentgruppe im Nachtest	132
Standardabweichung der Treatmentgruppe im Nachtest	42
Arithmetisches Mittel der Kontrollgruppe im Nachtest	88
Standardabweichung der Kontrollgruppe im Nachtest	86
KALKULATION DER EFFEKTGRÖSSE	
Effektgröße	259
Prozentwert kooperativer Reaktionen	* 167
Metrik der Effektgröße	259

Die mit * gekennzeichneten Items werden in der Folge erläutert.

Explikation des Kodiersystems

Identifikation der Effekte und der Untersuchungen
Für die Datenverarbeitung und Zuordnung der Ergebnisse zu einzelnen Studien wurden die Untersuchungen alphanumerisch kodiert.

Tabelle 16: Übersicht der in die Meta-Analyse eingeflossenen Studien

Code.	No. in M-A	Autoren	Jahr der Publik.	Anzahl der Effektgrößen
0101	2	AZRIN & LINDSLEY	1956	1
0201	1	BRAVER & BARNETT	1976	9
0202	3	BUELL et al.	1968	2
0203	24	BROTSKY & THOMAS	1967	1
0204	30	BIXENSTINE et. al.	1963	4
0205	35	BRAVER & ROHRER	1978	2
0206	38	BASS	1976	4
0301	49	CARLTON et al.	1975	1
0601	39	FEDER	1978	4
0701	5	GUMPERT, DEUTSCH & EPSTEIN	1969	5
0702	6	GALLO & WINCHEL	1970	3
0703	25	GARNER & DEUTSCH	1974	12
0801	29	HOGAN, FISHER & MORRISON	1974	8
0803	32	HARRISON & McCLINTOCK	1965	4
0804	46	HOECKELE	1972	2
0901	42	IBELMO	1976	2
1001	7	JENSEN & MOORE	1977	1
1101	8	KIRBY & TOLER	1970	1
1201	9	LINDKOLD, BAHAGAN & TEDESCHI	1970	6
1202	13	LINDSLEY	1966	9
1203	36	LINDSKOLD & COLLINS	1978	8
1204	43	LOEB	1975	12
1301	10	MADSEN	1967	8
1302	11	MARWELL, SCHMITT & SHOTOLA	1971	9
1304	14	MITHAUG	1969	2
1305	16	McNEEL	1973	4
1306	33	McCLINTOCK & McNEEL	1966	8

Fortsetzung von Tab. 16

Code	No. in M-A	Autoren	Jahr der Publik.	Anzahl der Effektgrößen
1307	34	MINAS et al.	1960	12
1308	44	MARTIN	1974	18
1309	45	MOELLER	1974	2
1310	4	McCLINTOCK & McNEEL	1967	10
1401	41	NACSON	1976	2
1501	15	OSKAMP & PERLMAN	1965	11
1502	40	O'DONELL	1978	12
1601	17	PAULSON	1974	3
1801	18	ROSENBERG	1960	1
1901	19	SLABY & CROWLY	1977	7
1902	20	SERBIN, TONICK & STERNGLANZ	1977	3
1903	21	SHAW	1958	2
1904	50	SCHNEIDER	1973	12
2001	22	TORRANCE	1971	2
2002	47	TURNER	1970	1
2003	48	THEROUX	1974	12
2201	23	VOGLER	1968a	1
2202	26	VOGLER	1968b	1
2203	27	VOGLER, MASTERS & MORILL	1970	1
2204	28	VOGLER, MASTERS & MORILL	1971	2
2301	37	WILSON	1971	12

Publikationsmerkmale

Für Meta-Analysen wurde inzwischen die Tradition etabliert, Merkmale zur Publikation zu kodieren. So lassen sich durch entsprechende Blockungen sowohl Verzerrungen wie auch eventuelle Entwicklungen ablesen. Innerhalb eines Forschungsprogramms wäre zu erwarten, daß über die Zeit Effekte größer werden oder die Varianz der Effekte abnähme, d.h., daß die vermuteten Prädiktoren zunehmend besser extrahiert worden wären.

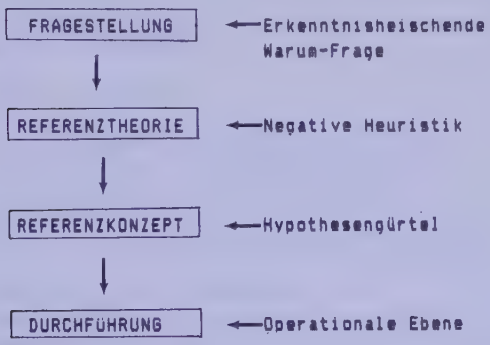
Anzahl der Effekte pro Studie

Tabelle 16 zeigt, daß die kleinste Analyseeinheit für die Integration nicht die Studien bilden, sondern die Effekte. In der vorliegenden Synthese streuen diese zwischen 1 und 18 pro Studie.

Referenztheorie, Fragestellung, Referenzkonzept und Problemstellung

WOLF (1973) unterscheidet in seinem Beitrag zur kritischen Reflexion empirischer Untersuchungen auf der vertikalen Ebene zwischen Relevanz der Untersuchung', 'Fragestellung und Hypothesen der Untersuchung' und Forschungsstrategie'. Fur das vorliegende Kodiersystem wurde dieser Gedanke aufgegriffen. Untersuchungen versuchen, WARUM-Fragen mit dem Ziel, den Informationsgehalt einer Theorie - auch unter Integration von Moderatorvariablen - zu erhöhen und/oder Theorien - auch im Hinblick auf ihre Hintergrundannahmen - zu bewähren, zu beantworten. Dabei kann nach LAKATOS mit Hilfe der negativen Heuristik von einem wahren' Theoriekern ausgegangen werden. Empirische Untersuchungen sollten - eine kritisch-rationale Wissenschaftstheorie unterstellt - daher auf eine Referentheorie verweisen. Jede Theorie umfaßt jedoch unterschiedliche Terme bzw. Konstrukte, mit deren Hilfe Phänomene erklärt werden sollen. Für die jeweilige Untersuchung muß damit im nächsten Schritt aus dem Theoriekern ein Referenzkonstrukt deduziert werden. Aus ihm ist dann (idealiter) die konkrete Vorgehensweise begründbar:

Abb. 11: Theoretische Konsistenz einer Untersuchung



Die Kodieritems Fragestellung, Referenztheorie und Referenzkonzept wurden in das Analyseschema aufgenommen, um

(a) die Erklärung von Kooperation vor dem Hintergrund ihrer Theorien aufzuzeigen
(b) Variablen-Aggregate, d.h. unter denselben Konstrukten zusamgefaßte Variablen, vergleichen zu können - ungeachtet der

Tabelle 18: Kategorien zur Fragestellung

---

01 KONTINGENTE BEKRAFTIGUNG
Hierunter fallen sowohl Fragen, welche Kontingenzen wirken, ob auch soziale Belohnung einen Effekt erzielt, ob Gruppenverstärkung gegenuber Individualbekräftigung eine hinreichende Bedingung darstellt, welche Revenuen eine Zunahme kooperativen Verhaltens bewirken, ob die Kenntnis der Kontingenzen zur Vergrößerung des Effekts beiträgt

02 MODELLERNEN
Untersuchungen dieser theoretischen Provenienz unterscheiden zwischen Fragen nach der Wirkung der Reaktionsdarbietung gegenüber der Kontingenzdarbietung. Weiter wird die Modellierbarkeit von Spiel-Strategien (Kontingenzen der Aktionsmuster) hinterfragt

03 VARIATION DER MATRIX-WERTE
Die Forschung zum (kooperativen) Wahl-Verhalten im PDG diskutiert ausführlich die Wirkung der Zellenindices auf die Rate kooperativer Wahlen. Die damit verbundenen Fragestellungen wurden hier subsummiert

04 HINWEISE ZUM KO-AKTEUR
Welche Variablen wirken zu kooperativem Verhalten interferent? Verschiede Problemlösungen zur Interferenz werden als Hinweise zur Kooperativität des Mitspieler und/oder der Situation diskutiert, z.B. der Einfluß induzierter Gruppennormen oder die Rate kooperativen Verhaltens des simultanen Mitspielers

05 AUFGABENSTRUKTUR UND ENTWICKLUNG
Die Komplexität der Aufgabenstruktur beeinflußt die Komplexität der sozialen Interaktionen. Mit komplexen Aufgaben und Situationen umgehen zu können, ist altersabhängig. Hier wird insbesondere danach gefragt, ob die Transparenz oder die Einfachheit der Aufgabenstruktur für Kinder erleichternd wirkt und ob Motive und soziale Fertigkeiten antezendent (mit-)wirken

---

Tab. 19: Kategorien der Variablen-Konstrukte

01 KONTINGENTE VERSTÄRKUNG KOOPERATIVER REAKTIONEN
kontinuierliche Gruppenverstärkung
Vorerfahrung im Verstärkungsniveau
Revenuen der Belohnungseinheiten
kontinuierliche soziale Verstärkung
individuelle vs Gruppenverstärkung
individuelle Verstärkung kooperativer Episoden
kontinuierliche Gruppenverstärkung
Latenz der Verstärkung
differentielle Belohnung und Bestrafung
Instruktion über die Belohnungskontingenz
Diskrimination
kontingente soziale Bekräftigung prosozialen Verhaltens
02 MODELLERNEN
reale Modellierung der Verstärkungsstruktur
symbolische Modellierung der Verstärkungsstruktur
selbstinduzierte Verstärkung
Modelldarbeitung der Reaktionskontingenz durch Fernsehen
Modelldarbietung interferenter Verhaltensweisen
Modelldarbietung kooperativer Interaktionen
03 VERTRAUENSINDUKTION
Verläßlichkeit, mit der eine angekündigte kooperative
Reaktion ausgeführt wird
unkontingente Prozentrate des Konföderanden
Strategie des Ko-Akteurs
04 INDUZIERTES MOTIV
Variation der Matrixindices
Induktion von Gruppennormen
Gewinnorientierung
Instruktion
Zielorientierung
Attribution
05 RESTKATEGORIE
ländliche gegenüber städtischer Sozialisation
Alter der Teilnehmer
Aufgabenstruktur

Die unabhängigen Variablen

Während die vorhergehenden Kategorien die Vergleichbarkeit der Effektgrößen in Abhängigkeit des Treatments auf der Konstruktebene ermöglichen, wird mit den nachstehenden Items versucht, die unterschiedlichen Treatmentanwendungen auf der funktional-analytischen Ebene zu beschreiben. Hierzu wurden alle Treatmentdarstellungen zunächst herausgeschrieben und anschließend in funktional-analytischer Schreibweise widergegeben (vgl. Anhang V). Kooperatives Verhalten ist demnach eine Funktion des Umfangs der Bekräftigungskontingenzen und/oder diskriminativer Stimuli. Die Diskriminationen umfassen auch Modellverhalten, einschließlich seiner unterschiedlichen Kontingenzen, Instruktionen oder Aufforderungen. Diese Analysen wurden zu Kategorien verdichtet und je in eine Rangreihe gebracht. Dabei wird angenommen, daß die Effektgrößen eine monotone Funktion der Rangposition innerhalb eines Variablen-Sets sind. Auf eine explizite Formulierung entsprechender Hypothesen wurde dabei verzichtet. Tabelle 20 bis 25 weisen die Rationale der Rangpositionen aus.

Tabelle 20: Rangordnung der UV-Realisierung: Modellernen

Kategorie	Rang	Begrundung des Rangplatzes
Modelle zeigen einen hohen Anteil nicht-kooperativer Verhaltensweisen	1	Nach BANDURA wird solches Verhalten akquiriert und gespeichert, daß positive Konsequenzen nach sich zieht. Nicht-kooperatives Verhalten liefert diskriminative Hinweise dafür, mit anderen Verhaltensweisen ein Ziel erreichen zu können. Zu erwarten ist, daß Beobachter, die Modelle erleben, die nicht-kooperative Strategien anwenden, selbst niedriges Kooperationsverhalten zeigen.
Tit-for-Tat ohne Explikation der Konsequenzen UND Modellierung prosozialer Verhaltensweisen	2	Die Tit-for-Tat -Strategie ist eine schwierige Interaktionsmöglichkeit. Der Beobachter muß ohne Information die Regeln erkennen. Kooperation unterscheidet sich von anderen prosozialen Verhaltensweisen durch durch aktuelle Reziprozitätsregeln, die durch die Modellierung prosozialen Verhaltens nicht mitgeliefert werden.
Tit-for-Tat einschließlich der Explikation der Kontingenz	3	Informiert die Tit-for-Tat-Darstellung zusätzlich über die Regeln dieser Strategie, ist zu erwarten, daß die Beobachter eine höhere Kooperation zeigen.

Fortsetzung der Tabelle 20

Kategorie	Rang	Begründung des Rangplatzes
		Die Darstellung einer kooperativen Periode zeigt die reziproke Ausführung gebender Reaktionen mit der Folge, ein Ziel gemeinsam erreicht zu haben. Die Zielerreichung bildet eine hinreichende Kontingenz für kooperative Perioden.
individuelle Verstärkung kooperativer Perioden	4	Die zusätzliche, externe Verstärkung reziprok-gebender Reaktionen liefert dem Beobachter Hinweise auf die Kontingenz der kooperativen Interaktion.
individuelle Verstärkung kooperativer Perioden bei gleichzeitiger Bestrafung nicht-kooperativen Verhaltens	5	Es wird erwartet, daß die ergänzende Bestrafung nicht-kooperativen Verhaltens die Modellierung kooperativer Reaktionen erhöht.

Tabelle 21: Rangordnung der UV-Realisation: Bekräftigungskontingenz VOR dem Meßvorgang

Kategorie	Rang	Begründung des Rangplatzes
Belohnung prosozialen Verhaltens	1	Der Modellierungsbedingung ähnlich, liefert die Belohnung prosozialen Verhaltens keine Hinweise über die Struktur kooperativer Reaktions-Interdependenzen. Es ist daher eine niedrige Kooperationsrate zu erwarten.
Bekräftigung gebender Reaktionen	2	Werden unabhängig vom Verhalten des Ko-Akteurs gebende Reationen bekräftigt, wird Altruismus ausgeformt, nicht die Interdependenz gebender Reaktionen.
geringe Bekräftigung kooperativer Perioden	3	Erst die Bekräftigung gebender Reaktionen unter dem Aspekt der Reziprozität führt zur Ausformung von kooperativen Interaktionen. Zwischen geringer und hoher Verstärkung wurde angesichts der unterschiedlichen Niveaus der Revenuen geblockt.
hohe Bekräftigung kooperativer Perioden	4	
Gruppenkontingente Bekräftigung einer Periode	5	Die individuelle Bekräftigung reziproker gebender Reaktionen hebt auf den Beitrag des einzelnen Gruppenmitgliedes ab. Die Verstärkung beider (aller) Interakteure steht unter der Kontingenz des gemeinsamen Handelns. Es wird erwartet, daß solche Bedingungen den höchsten Kooperationseffekt erzeugen.

In Analogie zur Bekraftigung kooperativer Reaktionen ist zu erwarten, daß für die Bekräftigung alternativer, nicht-kooperativer Reaktionen gilt: Je umfangreicher alternative Reaktionen unter kontingenter Kontrolle stehen, desto weniger kooperative Reaktionen sind zu erwarten. Nach diesem Rational wurde die folgende Rangreihe gebildet.

Tabelle 22: Rangordnung der UV-Realisierung: Belohnung alternativer Reaktionen VOR dem Meßvorgang

Kategorie	Rang	Begründung der Rangposition
Belohnung alternativen Verhaltens bei möglicher Bestrafung von kooperativen Reaktionen	1	Nach dem Rational differentieller Verstärkung wird solches Verhalten gezeigt, daß unter positiver Kontingenz steht, unter Vermeidung solchen Verhaltens, das unter aversiver Kontrolle steht.
hohe Bestrafung kooperativen Verhaltens	2	Reaktionen, die unter der Kontrolle strafender Reize stehen, werden vermieden. Selbst kooperative Interaktionen, die unter der Belohnungskontingenz innerhalb der kooperativen Reziprozität stehen, werden dann nicht oder nur in geringem Umfang ausgeführt, wenn sie extern bestraft werden.
niedrige Bestrafung kooperativen Verhaltens	3	Zwischen hohen und niedrigen Bestrafungskontingenzen wurde aufgrund der vielschichtigen Bestrafungsniveaus geblockt.
Alternatives Verhalten wird höher belohnt als kooperatives Verhalten	4	Die Reaktionsdifferenzierung unterstellt, daß das Verhalten ausgeformt wird, das der höher ausfallenden positiven Konsequenz unterliegt
Kooperatives Verhalten wird höher belohnt als alternatives Verhalten	5	

Die UV "Diskriminative Stimuli VOR dem Meßvorgang (DHV)" und die UV "Diskriminative Stimuli WÄHREND des Meßvorgangs (DHW)"

Wie weiter oben aufgezeigt, sind diskriminative Stimuli kontingenzanzeigende Hinweise für bestimmte Reaktionen. Die Palette solcher Hinweise im Rahmen der Kooperationsforschung ist umfangreich. Die in den Untersuchungen gefundenen Stimuli zur Kooperativität der Situation und/oder der Person (des Ko-Akteurs) erlaubten es nicht, ein trennscharfes Kategoriensystem zu entwickeln. Daher wurde eine fünfpolige Ratingskala unterlegt, auf der der Hinweischarakter der einzelnen diskriminativen Informationen zugeordnet wurde:

Die Situation ist...

nicht			hoch
kooperativ	ambivalent		kooperativ

Ankerpunkte bildeten auf der einen Seite solche Stimuli, die entweder auf die Nicht-Kooperativität oder auf der anderen Seite auf die eindeutige Kooperativität der Situation hinwiesen. Mittelpunkt dieser Skala bilden solche Hinweise, die die Ambivalenz der Situation anzeigen.

Die UV "Belohnungskontingenzen während des Meßvorgangs (KBW)"

In Analogie zu den Variablen KBV und KBA wurde für die Kontingenzstruktur während des Meßvorgangs eine separate Variable entwickelt. Mit ihr sind "Belohnung kooperativer Reaktionen", "Belohnung alternativer Reaktionen" und "Bestrafung kooperativer Reaktionen" in einer Rangreihe zusammengefaßt.

Tabelle 23: Rangordnung der UV-Realisation: Kontingenzen während des Meßvorgangs (KBW)

Rangplatz	Inhalt dieser Rangposition
1	gebende Reaktionen können bestraft werden
2	gebende Reaktionen werden geringer belohnt als alternative Reaktionen
3	kooperative Reaktionen werden höher belohnt als alternative Reaktionen
4	individuelle Bekräftigung kooperativer Episoden
5	individuelle Bekräftigung kooperativer Reaktionen unabhängig vom Verhalten des Ko-Akteurs

Tit for Tat (deutsch: Wie Du mir, so ich Dir)

Tit-for-Tat (TFT) ist eine kontingente Verhaltensstrategie im PDG, von der angenommen wird, die Kooperationsrate des Akteurs wirksam erhöhen zu können. Der TFT-Spieler beginnt nie, von sich aus eine unkooperative Strategie zu spielen (er 'mogelt' nicht ohne Grund). Spielt der Akteur unkooperativ, reagiert TFT unmittelbar und einmalig mit Vergeltung, d.h. TFT mogelt im nächsten Zug, ohne jedoch nachtragend zu sein. Diese Strategie kehrt sofort wieder zu kooperativen Wahlen zurück, wenn der Akteur auch kooperiert. TFT-Varianten, z.B. 90% TFT plus 10% Mogelreaktionen, wurden nicht als TFT-Strategie kodiert.

Prozentaler Anteil unkontingenter kooperativer Reaktionen

Im Rahmen der PDG-Forschung wird danach gefragt, ob die Höhe der - unkontingenten - kooperativen Wahlen des Simultanspielers die Kooperationsbereitschaft der Akteure beeinflussen. Eine solche, vorher festgelegte Reaktionsquote liegt vor, wenn der Simultanspieler unabhängig vom Wahl-Verhalten der Probanden diese fixe Reaktionsquote im Spielverlauf äußert. In der Literatur wird ein positiver Zusammenhang zwischen Quotenhöhe und kooperativen Wahlen der Probanden angenommen. Für die Analyse wurde der prozentale Anteil kooperativer Wahlen des Ko-Aktuers (Simultanspielers) kodiert.

Wiedergabe der Spielzüge (Display)

HAKE und Mitarbeiter (HAKE 1973, HAKE & VUKELICH 1973) untersuchen mit Verweis auf FESTINGER's Annahme des sozialen Vergleichs (FESTINGER 1954) die Wirkung der Reaktions- und/oder Ergebnisanzeige auf den Kooperationsverlauf. Sie unterscheiden dabei zwischen 'self-auditing', d.h. die Anzeige des eigenen Verhaltens bzw. seiner Konsequenzen und 'coactor auditing' als die Anzeige des Verhaltens bzw. der Ergebnisse des Ko-Akteurs. Diese Widerspiegelung ermöglicht, das eigene Verhalten hinsichtlich der Kosten und des Nutzens in Relation zum Verhalten des Interaktionspartners zu setzen (vgl. auch HAKE & OLVERA, 1978). HAKE et al. (1973) zeigen in einem Experiment, daß Probanden im Mittel nach jeweils zwei Perioden von dieser Informationsquelle Gebrauch machen. In der Interpretation lerntheoretischer Paradigmen bilden die Wiedergabe des eigenen Ergebnisses und das des Partners sowohl diskriminative Hinweise (welcher Teilnehmer die effektivere Strategie spielt und/oder ob eine kooperative Interaktion vorliegt) sowie eine Bekräftigungsquelle. Gleiche Punktwerte halten wahrscheinlicher die reziproke Interaktion aufrecht. Punktabweichungen liefern sowohl Hinweise auf eine fehlende Reziprozität wie sie bei einer negativen Differenz die eigene Strategie bestrafen, bzw. bei positiver Abweichung die eigene Strategie belohnen. Verschiedene Autoren nehmen an, daß die Wiedergabe des gegenseitigen Spielverhaltens eine unmittelbare Vergeltung bei Verletzung der Reziprozität erzeugt und dies zur weiteren Abnahme der Reziprozität führt.

Revenuen

In den Experimenten wird das kooperative und/oder alternative Verhalten vor und während des Meßvorgangs durch verschiedene Arten der Verstärker auszuformen bzw. zu variieren versucht. Verschiedene Autoren betonen, daß im Spiel erhaltene Punkte als Form symbolischer Bekräftigung keine hinreichende Bedingung für die Erhöhung der Auftrittswahrscheinlichkeit kooperativen Verhaltens darstellen. Im Rahmen der PDG-Forschung wird diskutiert, ob Punkten oder Spielgeld überhaupt der Anreiz zukomat, interferente Faktoren zu kompensieren. Aus diesem Grund wurde verschiedentlich der Eintauschwert (Revenuen) variiert.

Tabelle 24: Rangordnung der Eintauschwerte

Rangplatz	Eintauschwert
1	Punkte
2	imaginäres Geld (niedrig)
3	imaginäres Geld (hoch) Süßigkeiten (für Kinder) soziale Verstärkung
4	echtes Geld (niedrig)
5	echtes Geld (hoch)

Matrixwerte

Nullsummen- wie Nichtnullsummenspiele haben sich aus der Utilitarismustheorie der ökonomie und aus der mathematischen Spieltheorie entwickelt. Solche Spiele sind zunächst nur interaktionsbedingt und aus der Tradition heraus nicht zwingend auf eine Theorie zuruckzuführen. So verzichten viele Autoren auf explizit theoretische Basisannahmen und versuchen, den Kooperationswert durch spielimmanente Bedingungen zu variieren. Eine Menge von Variablen bilden dabei die Matrixwerte. Nach RAPOPORT (1968) existieren für 2x2-Spiele (d.h. zwei Spieler mit jeweils zwei Verhaltensalternativen) acht unabhängige Variablen, für die bei drei Ausprägungsgraden (niedrig, mittel, hoch) 3**8 = 6561 Experimente notwendig wären, um den Einfluß der Matrixwerte auf das Wahlverhalten näher zu bestimmen. Für die hier integrierten Kooperationsuntersuchungen sind die Zellenwerte fur die Ko-Akteure jedoch identisch. Die Matrixwerte REward, Temptation, Sucker's payoff und Punishment wurden ebenfalls kodiert.

Instruktion

Verschiedene Arbeiten unterstellen den Einfluß der Instruktion zum Spielverhalten auf das Wahlverhalten im PDG. Für die Analyse wird angenommen, daß in folgender Rangordnung dieser Einfluß wächst:

Tabelle 25: Zunehmender Einfluß der Instruktion auf kooperative Wahlen

Rangposition	Inhalt der Instruktion
1	mache einen eigenen Gewinn
2	dies ist ein Entscheidungsspiel
3	Kooperation wie Kompetition sind mögliche Reaktionen
4	(01) + explizite Bekanntgabe der Spielrationale
5	(02) + explizite Bekanntgabe der Spielrationale
6	(03) + explizite Bekanntgabe der Spielrationale

Klassifikation der abhängigen Variablen

Den wenigen nominalen Definitionen stehen eine Fülle operationaler Definitionen von Kooperation in Form unterschiedlicher Meßinstrumente zur Verfügung. Ihr Einfluß auf das Kooperationsverhalten wird umfangreich diskutiert. HAKE & VUKELICH (1972) unterscheiden zwischen Handlungs-(performance) und Wahl- (choice) Verfahren.

"In a performance procedure the rates of cooperative responding under the different values of a variable are compared by studying the values one at a time. A choice procedure, however, is analogous to a concurrent situation, in which two or more values of a variable are studied at the same time by measuring the extent to which each value is chosen" (S. 221).

Für die Kodierung der verschiedenen Meßinstrumente wurden neben dieser Unterscheidung die weitere Differenzierung von Handlungsinstrumenten nach HAKE & OLVERA (1978) aufgegriffen. Sie klassifizieren kooperative Interaktionen auf 5 Dimensionen:

INTERDEPENDENTE VERSUS DEPENDENTE KOOPERATIONSKONTINGENZEN
In kooperativen Kontingenzen hängt die Verstärkung der einen Person zumindest zum Teil vom Verhalten der anderen Person ab. Kooperationskontingenzen sind stark dependent, wenn der Verstärker des Akteurs vollkommen unter der Kontrolle des Verhaltens des Ko-Aktuers steht. Schwache Dependenz liegt

vor, wenn die Verstärkung stark vom eigenen Verhalten und nur schwach vom Verhalten des Ko-Akteurs abhängig ist:

$$m * (A)R(G) \longrightarrow n * (B)R(I) \longrightarrow (B)K+$$

m gt n starke Dependenz
m lt n schwache Dependenz

Kooperative Kontingenzen sind interdependent, wenn das Verhalten beider Akteure von einem Verstärker abhängig ist:

$$(A)R(G) + (B)R(G) + \ldots\ldots\ldots\ldots \longrightarrow (A,B)K+$$

REAKTIONSTEILUNG (RESPONSE SHARING) VERSUS REAKTIONSAUSTAUSCH (RESPONSE EXCHANGE)
Reziprozität ist innerhalb einer jeden Episode oder aber über eine (lange) Periode hinweg möglich. Reziprozität innerhalb einer jeden Episode bezeichnen die Autoren als Reaktionsteilung, über eine Periode hinweg als Reaktionsaustausch.

Beispiel
Wenn für ein Nicht-Nullsummen-Spiel die Bedingung 2Re = (T + S) gilt (was vom Konfliktpotential her uninteressant ist) kann Kooperativität sowohl über Reaktionsteilung (a1b1) + (a1b1) + ... wie auch über Reaktionsaustausch aufrechterhalten werden (a1b2) + (a2b1) + (a1b2) + (a2b1) + ...

KOSTEN VERSUS KOSTENFREIE KONTINGENZEN
Führt eine kooperative Reaktion zur eigenen Verstärkung und zu der des Ko-Akteurs, enstehen dem Akteur keine Kosten. Umgekehrt erzeugen solche gebenden Reaktionen Kosten, auf die keine eigene Verstärkung erfolgt:

$$(A)R(G) + (B)R(I) \longrightarrow (B)K+$$ Kosten für A
$$(A)R(G) + \ldots \longrightarrow (B)K+ + (A)K+$$ Kostenfreie Kontingenz für A

FORCIERTE KOOPERATION VERSUS ALTERNATIVE REAKTIONEN
Interaktive Situationen unterscheiden sich in dem Grad, mit dem sie alternative Reaktionen verstärken. Werden nur kooperative Reaktionen bekräftigt, sprechen HAKE & OLVERA (1978) von forcierter Kooperation. Demgegenüber existieren solche Kontingenzen, die alternative Verhaltensweisen ebenso belohnen.

NICHTSOZIALE VERSUS SOZIALE SITUATIONEN
Der Ko-Akteur liefert diskriminative Stimuli, die unter dem Begriff der sozialen Emergenz (d.i. der Einfluß auf Verhalten durch die bloße Anwesenheit anderer Personen) diskutiert werden (vgl. LINDSLEY 1966). Um diesen Einfluß kontrollieren zu können, wurden Experimente durchgeführt, in denen Akteure kooperierten, ohne irgendwelche Informationen über die Existenz

eines Ko-Akteurs zu erhalten (vgl. SIDOWSKI 1957, LINDSLEY 1966). Solche Situationen werden als nicht-soziale Situationen bezeichnet. Auf der anderen Seite nehmen die Interaktionspartner sich in unterschiedlichem Maße wahr. Die folgende Tabelle bildet die Kodierung der Meßinstrument entlang der Differenzierungsmöglichkeiten ab.

Tab. 26: Klassifikation der Meßinstrumente

Kodier-nummer	Typ
1	IND-SHAR-NONS-FORCE
2	IND-SHAR-NONS-ALT
3	IND-SHAR-S-FORCE
4	IND-SHAR-S-ALT
5	IND-EXCH-NONS-FORCE
6	IND-EXCH-NONS-ALT
7	IND-EXCH-S-FORCE
8	IND-EXCH-S-ALT
9	DEP-EXCH-NONS-FORCE
10	DEP-EXCH-NONS-ALT
11	DEP-EXCH-S-FORCE
12	DEP-EXCH-S-ALT
14	PDG *
20	BEBACHTUNG
25	SONSTIGES

IND = independent; DEP = dependent; SHAR = Reaktionsteilung; EXCH = Reaktionsaustausch; NONS = nichtsoziale Situation; S = soziale Situation; FORCE = forcierte Kooperation; ALT = alternative Reaktionen werden bekräftigt
* Für PDG-Experimente wurden die Zellenwerte kodiert

Validität der Studien

Der Einfluß der Studienvalidität auf die Güte und den Umfang des Effekts steht im Mittelpunkt der kritischen Diskussion um die Meta-Analyse. Um eventuelle Verzerrungen durch Validität zu kontrollieren, wurde in Analogie zu COOK & CAMPBELL (1979) zwischen innerer, äußerer, Konstruktvalidität und Validität des statistischen Schlusses unterschieden. Die innere Gültigkeit informiert darüber, wie valide aus der Untersuchung gefolgert werden kann, daß die untersuchte Beziehung auch tatsächlich auf den von dem Experimentator unterstellten Zusammenhang zurückführbar ist. Die externe Gültigkeit prüft demgegenüber die Generalisierbarkeit dieser Beziehung auf andere als die untersuchten Meßmethoden, Personen, Bedingungen und Zeitpunkten. Die Validität des statistischen Schlusses ermöglicht zu beurteilen, ob der

festgestellten Zusammenhang mit den Daten überhaupt vereinbar ist. COOK & CAMPBELL (1979) nennen drei Entscheidungen, die bei der Evaluation von Experimenten gefällt werden müssen:

"(1) Is the study sensitive enough to permit a reasonable statement about covariation?

(2) If it is sensitive enough, is there any reasonable evidence from which to infere that the presumed cause and effect covary? and

(3) if there is such a evidence, how strongly do the two variables covary?" (S. 39).

Mit der Konstruktvalidität wird die Übereinstimmung zwischen Operation und Konstrukt überprüft. "Can I generalize from this one operation or set of operations to a referent construct?" (COOK & CAMPBELL, 1979, S. 39).

Die vier Gültigkeitsaspekte wurden mit Hilfe einer dreipoligen Skala (niedrig, mittel, hoch) für jede Effektgröße geschätzt.

Kalkulation der Effektgröße

Teil I stellte die unterschiedlichen Möglichkeiten zur Schätzung der Effektgröße vor. Um Verzerrungen durch die unterschiedlichen Schätzmodalitäten kontrollieren zu können, wurde die Datenbasis, auf deren Grundlage der Effekt geschätzt wurde, wie folgt kodiert.

Tabelle 27: Datenbasis für die Effektgößen

Kode	Datenbasis
1	Prozentwerte der Treatmentgruppe
2	Nichtparametrische Statistiken
3	Mittelwerte und Standardabweichungen
4	Teststatistiken (F, t)
5	Statistiken aus Einweg-Varianzanalyse
6	Statistiken aus der mehrfaktoriellen VA
7	Korrelationen
8	empirische p-Werte
9	kritische p-Werte
10	Sonstiges

Tabelle 28: Berechnungsmodalität der Effektgrößen

Kode	Datenbasis
1	Mittelwertdifferenz und Standard-abweichung der Kontrollgruppe
2	Mittelwertdifferenz und Schätzung der Standardabweichung durch MS(w)
3	Mittelwertdifferenz und Schätzung der Standardabweichung aus MS(tot) - MS(treatment)
4	Probit
5	Chi**2
6	t-Tabelle
7	AM(t) und empirischer p-Wert
8	AM(t) und kritischer p-Wert
9	nicht-parametrische Werte
10	Korrelationen
11	Schätzung aus F bei mehrfaktorieller VA
12	Sonstiges

Tabelle 29: Quelle des Mittelwertes

Kode	Quelle
1	unkorrigierter Posttest-Wert
2	Kovarianzkorrigierter Wert
3	Residualzuwachs
4	Vortest-Nachtest-Differenz
5	Restkategorie

Kalkulation der Effektgröße bei Prozentwerten kooperativen Verhaltens

Soweit die publizierten Statistiken eine Effektberechnung bzw. -schätzung erlaubten, wurden die von GLASS, McGAW & SMITH (1981) vorgeschlagenen Kalkulationen eingesetzt. Eine Vielzahl von Experimenten im Rahmen der PDG-Forschung gibt im Regelfall jedoch nur die 'Rate kooperativer Reaktionen' in Prozentwerten an, bzw. bildet sie graphisch ab. Somit fehlen die für die Berechnung der Effektgröße notwendigen Referenzwerte der Kontrollgruppe. Für diese Fälle wurde ein Modell entwickelt, das die Effektschätzung über

die Probit-Analyse erlaubt (Abschnitt 1.4.5.). Ohne Einfluß irgendwelcher Bedingungen auf das Wahlverhalten eines Probanden gibt es beim PDG vier alternative Wahlmöglichkeiten:

(i)	eine kooperative Wahl, um zu kooperieren	(R(G) - R(G))
(ii)	eine kooperative Wahl, um sich ausbeuten zu lassen	(R(G) - R(N))
(iii)	eine kompetitive Wahl, um auszubeuten	(R(N) - R(G))
(iv)	eine kompetitive Wahl in Erwartung einer solchen durch den Ko-Akteur	(R(N) - R(N))

Dieses naive Modell erlaubt die Bestimmung eines pc-Wertes von pc = .25, dem ein Prozentwert kooperativer Wahlen von p(e) = (%R(G)/100) gegenüber gestellt wird. Mit Hilfe der Tabelle aus GLASS, McGAW, & SMITH, (1981, S. 139) können die entsprechenden Effektgrößen bestimmt werden. Die geringe Validität dieses Modells wird bewußt in Kauf genommen. Ziel der vorliegenden Meta-Analyse ist nicht die Extraktion handlungsrelevanter Effektgrößen, sondern die Bestimmung wirksamer Einflußgrößen auf kooperative Reaktionen. Die vorliegende Schätzung liefert eine lineare Transformation von Prozentwerten in Effektgrößen. Bei Vorlage eines anderen Modells würde jedoch die Rangreihe der Effekte erhalten bleiben.

Prozentwert kooperativer Wahlen

Neben dem Effekt wurde als abhängige Variable auch der Prozentwert kooperativer Wahlen kodiert. Bei Vorlage von Graphiken, die die Spielzüge zusammenfaßten, wurde mit Hilfe von Millimeterpapier der Prozentwert abgelesen und ungewichtet gemittet. Eine separate Analyse erbrachte keine Hinweise auf systematische Veränderungen im Kooperationsverhalten über den Spielverlauf.

5. Ergebnisse

5.1 Deskription

Im vorliegenden Teil bilden die Ergebnisse auf deskriptiver Ebene die Datenunterschiede im entwickelten Kodiersystem ab. Eine Vielzahl von Ergebnissen geht auf multiple Variableneinflüsse zurück. Gemittelte Effekte ermöglichen daher keine Aussage, daß die jeweiligen Effekte allein auf die geblockten Variablen zurückzuführen sind. Solange unterschiedliche zusätzlich Einflußquellen vorliegen, liefern Effekt-Mittelwerte nur Hinweise über den Einfluß der geblockten Variable. Plausible , konfundierte Bedingungen werden in der Folge deshalb mit dargestellt.

In die Meta-Analyse flossen insgesamt 49 Individualstudien mit n = 259 Effektgrößen ein. Der mittlere Effekt liegt bei AM(ES) = .652 mit einer Standardabweichung von SD(ES) = .759 (Standardfehler = .047). Abbildung 12 weist sowohl Effektgrößen unterhalb der Kontrollgruppe (lt 0) aus, wie auf der anderen Seite verschiedene Experimente zu überdurchschnittlich hohen Effekten führen. Im globalen Vergleich zur Kontrollgruppe verhält sich ein Teilnehmer bei einem Experiment, in dem Kooperation ausgeformt wurde, im Mittel kooperativer als 74% der Mitglieder der Kontrollgruppe.

Abb. 12: Stem-and-leaf-Darstellung der Effektgrößen (Faktor = 0.1)

```
-15  3
-14
-13
-12  7
-11
-10  9
-09
-08  5
-07  16
-06  6
-05  014
-04  02
-03  047799
-02  224579
-01  123447777
-00  0000033389
```

```
+00  23344558999
+01  1234555556666779999
+02  023337778899
+03  111124556666899
+04  1222334566689
+05  1112334444556677789    **********  Die einzelnen Effektgrößen
+06  000122234557777899                 können wie folgt am Beispiel
+07  0000125555678                      der Zeile (****) gelesen werden:
+08  00112244678                        für +05 ( * 0.1) = +0.51: 0.51:
+09  00222255888999                     0.51: 0.52: 0.53:...:0.58: 0.59
+10  122334466679
+11  14689999
+12  122345589
+13  0123349
+14  001389
+15  1779
+16  2459
+17  199
+18  01689
+19  035
+20  99
+21
+22  4
+23  11
+24
+25  8
+26
+27
+28
+29  7
+30
+31
+32
+33
+34  5
+35
+36
+37  5
+38
+39
+40  3
```

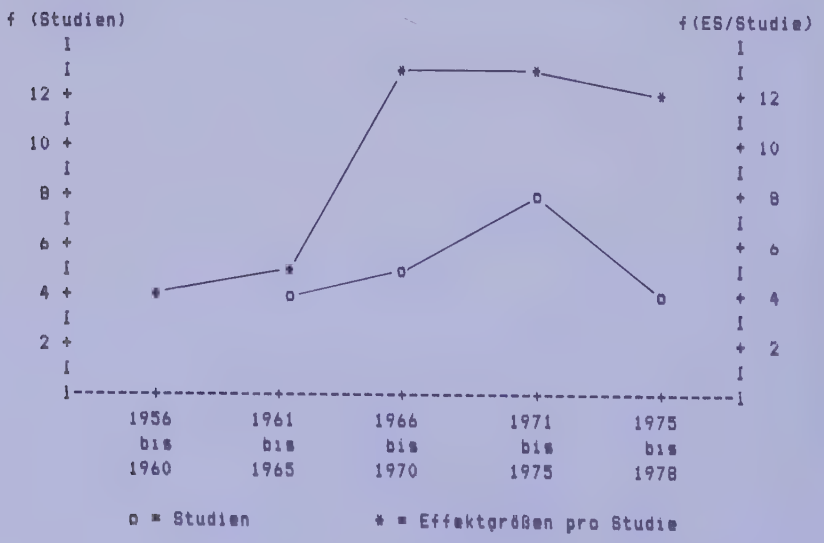
Die in die Analyse einfließenden Studien wurden zwischen 1956 und 1978 veröffentlicht.

Tabelle 30 Effektgrößen in Abhängigkeit des Publikationszeitraums

Publikationszeitraum	Effektgrößen AM	SD	SE	MD	N Stud.	N ES
1956 - 1960	.40	.46	.11	.37	4	16
1961 - 1965	1.01	.65	.15	1.22	5	19
1966 - 1970	.89	.58	.08	.71	13	59
1971 - 1975	.54	.69	.07	.46	14	111
1976 - 1978	.58	1.04	.14	.42	12	54

Abbildung 13 zeigt deutlich einen Anstieg in der Anzahl der Veröffentlichungen zwischen 1965 und 1970, sowie zwischen 1975 und 1978.

Abbildung 13 Veröffentlichungsentwicklung im Erfassungszeitraum



Diese Entwicklung weist auf ein zunehmendes Interesse an Fragestellungen zur Kooperationsforschung und entspricht dem Trend in der Erforschung prosozialer Verhaltensweisen (vgl. STAUB 1978). Bedeutende Arbeiten zur Akquisition prosozialer Verhaltensweisen unter verschiedenen Antezendenzen erschienen in diesem Zeitraum (BANDURAS Publikationen zwischen 1969 und 1977; BERKOWITZ 1962 bis 1970; BRYANs Beiträge zum altruistischen Verhalten 1970 bis 1975; CAMPBELLs sozialevolutionären Beiträge 1965, 1972; DEUTSCHs nachfolgenden sozialpsychologischen Studien zum Verhältnis von Kooperation und Vertrauen; LATANEs Theorie des sozialen Einflusses (1976) auf der Grundlage umfangreicher Untersuchungen zum 'Bystander-Effekt'). Diese Entwicklung mag durch eine Vielzahl von Ursachen erklärbar sein. Einige von ihnen fallen jedoch in diesen Zeitraum:

(a) das zunehmende Bewußtsein unsozialer Tendenzen in (amerikanischen) Gesellschaften

Beispiel
Verschiedene Beiträge zu prosozialem Verhalten verweisen auf den Fall Kitty Genoveses. Diese junge Frau wurde 1964 am frühen Abend nach einem langen Kampf durch eine Vielzahl von Messerstichen getötet, obgleich mindestens 30 Bewohner anliegender Häuser diesem Kampf zusahen. Die nicht geleistete Hilfestellung löste in den USA engagierte Diskussionen und in Folge dessen Untersuchungen zur Hilfsbereitschaft aus (vgl. STAUB 1978, THOMPSON 1980).

(b) die Diskussion zur Frage, ob prosoziale Verhaltensweisen prophylaktisch gegen konstatierte Aggressivität unter Jugendlichen implementierbar ist

(c) die Studentenunruhen und der damit verbundene Zeitgeist einer selbstbestimmten, prosozialen Gesellschaft

(d) die Entwicklung, internationalen Konflikten durch hohe Interdependenzen entgegenzuwirken.

Für diesen Zeitraum spiegeln die Ergebnisse keine Entwicklung in der Zunahme der Effektgrößen wider. Dies widerspricht der Vermutung, im Rahmen von Forschungsprozessen varianzerzeugende Variablen zunehmend kontrollieren zu können, d. h. auch, mit Forschung das Zielverhalten 'Kooperation' gesamtgesellschaftlich oder zumindest in wissenschaftlichen Untersuchungen positiv beeinflussen zu können.

Tabelle 31 Tau-Korrelationen zwischen Treatment-Bedingungen und Publikationsjahr

Variable	Tau	Signifikanz niveau
Modellernen	.18	.091
Kontingente Belohnung vor dem Meßvorgang	-	ns
Belohnung alternativer Verhaltensweisen	-.22	.113
Diskriminative Hinweise vor dem Meßvorgang	-	ns
Kontingente Belohnung während des Meßvorgangs	.21	.001
Instruktion zum Treatment	.22	.003
Versuchung (T)	-.35	.001
Treatmentdauer in Minuten	-.10	ns
Treatmentdauer in Spielzügen	-.35	.001

Die Tabelle zeigt, daß zwar zunehmend komplexere, d.h. im Sinne der Rangpositionierungen effektivere, Treatments eingesetzt wurden, deren Einfluß auf Kooperation unterstellt werden kann, daß aber auf der anderen Seite auch solche Treatments verstärkt appliziert wurden, die kooperatives Verhalten vermutlich unterminieren (zunehmende Belohnung alternativer Reaktionen, Erhöhung des Versuchungswertes (T)). In der PDG-Forschung wurde zudem im Forschungsprozeß die Anzahl der Spielzüge reduziert, obwohl die Frage unbeantwortet blieb, ob lange Spielfrequenzen nicht den Lernprozeß ermöglichen, Kooperation sei die optimalste Strategie.

Die experimentelle Kooperationsforschung wird durch Untersuchungen eingeleitet, die sich an der Theorie von DEUTSCH oder an der Theorie der operanten Konditionierung orientieren.

Tabelle 32 Orientierung an unterschiedlichen Referenztheorien im Verlauf des Erfassungszeitraumes

REFT	1956 - 60	1961 - 65	1966 - 70	1971 - 75	1976 - 78
OK	2	22		37	12
AM(ES)	1.13	.87		.60	1.15
SKLT	0	0	1	16	21
AM(ES)	-	-	2.58	.31	.40
AT	0	11	8	4	0
AM(ES)	-	1.39	.52	.16	-
FTD	14	4	26	50	20
AM(ES)	.3	.18	1.0	.58	.43

REFT = Referenztheorie; OK = Paradigma der operanten Konditionierung; SKLT = Sozial-kognitive Lerntheorie; AT = Austauschtheorie; FTD Feldtheorie der Kooperation sensu DEUTSCH

Die Tradition der Kooperationsforschung entspringt und steht in der Referenz von DEUTSCH, ohne jedoch mit Hilfe dieser Paradigmen hohe Effekte erzeugen zu können. Untersuchungen auf der Basis der operanten Konditionierung gelangen zu durchweg mittleren Effekten. Erst recht spät orientieren sich Studien an der Sozial-kognitiven Lerntheorie nach BANDURA.

DIE MITTLEREN EFFEKTGRÖSSEN IN ABHÄNGIGKEIT DER PUBLIKATIONSART

Weiter vorne wird die Verzerrung durch den Publikationstyp bereits diskutiert. Tabelle 33 zeigt unterschiedliche mittlere Effektgrößen.

Tabelle 33 Mittlere Effektgrößen in Abhängigkeit des Publikationstyps

Publikationstyp	AM(ES)	SD	SE	MD	N Stud	N ES	N/ Stud
Zeitschrift	.74	.66	.05	.7	36	177	4.9
Buch	.71	.49	.15	.55	1	11	11
Dissertation	.43	.96	.11	.27	11	71	6.5

Die Ergebnisse stimmen mit denen von M.L. SMITH (1980) überein. Sie zeigt folgende Verzerrungen auf Grund der Publikationsart auf:

Tabelle 34 Mittlere Effektwerte in Abhängigkeit der Publikationsquelle über mehrere Meta-Analysen nach M.L. SMITH (1980)

Publikationstyp	AM(ES)	N(ES)
Zeitschrift	.64	2204
Buch	.30	219
Dissertationen	.48	956

Beiträge in Dissertationen und Zeitschriften unterscheiden sich nicht nur im mittleren Effekt, sondern auch im Schwerpunkt des theoretischen Interesses.

Tabelle 35:
Referenztheorien in Abhängigkeit des Publikationstyps

Publikationstyp		OP	SKLT	AT	FTD	Rest
Journal	N	42	15	23	90	1
	%	24	9	13	51	-
Dissertation	N	20	23	-	24	2
	%	3	28	-	34	3

Dissertationen beziehen sich in ihren Untersuchungen im Rahmen der Kooperationsforschung im wesentlichen auf die sozial-kognitive Lerntheorie und auf die Theorie von DEUTSCH. Demgegenüber versuchen Zeitschriften-Beitrage die Wirkung von Bekräftigungskontingenzen auf die Ausformung von Kooperation zu untersuchen. Als unabhängige Variable wird in Zeitschriften daneben die Vetrauenskomponente untersucht, während in Dissertationen die Modellierung von Kooperationskontingenzen gegenüber der Modellierung von gebenden Reaktionen bevorzugt wird.

Tabelle 36 Untersuchtes Konstrukt in Abhängigkeit der Publikationsquelle

Publikationstyp		KV	DV	MKV	MRG	VT	MI	Rest
Zeitschrift	N(ES)	60	-	14	1	50	32	15
	%	34	-	8	-	28	18	9
Dissertation	N(ES)	-	2	23	18	12	12	4
	%	-	3	32	25	12	12	6

KV = Kontingente Verstärkung; DV = Differentielle Verstarkung;
MKV = Modellierung der Kooperationskontingenzen; MRG = Modellierung

der kooperativen Reaktionen: VT = Vertrauen; MI = Motivinduktion

In beiden Publikationstypen wird der Einfluß der Verstärkungskontingenz als wesentliche Einflußquelle unterstellt. Während Zeitschriften hier mit einem sparsameren Modell operieren, wenden Dissertationen ein elaborierters Modell an (und erzeugen niedrigere Effekte). Ein häufig untersuchtes Konstrukt in Zeitschriften bildet im Rahmen der Kooperationstheorie von DEUTSCH das Vertrauen. Die Frage, welche vertrauensbildenden Maßnahmen vor oder während des Meßvorgangs wirksam sind, stehen dabei im Mittelpunkt.

Tabelle 37: Tau-Korrelationen zwischen Publikationsart und verschiedenen kodierten Variablen

Publikationsart mit	tau	N(ES)	alpha
(A) Variablen zur Konsistenz der Untersuchung			
Explikation der UV	.33	259	.001
Operationalisierung der UV	.19	259	.001
Definition der AV	.5	259	.001
Transparenz der Fragestellung	.43	259	.001
(B) Variablen zur Validität			
Innere Gültigkeit	.24	259	.001
Äußere Gültigkeit	-.16	259	.005
(C) Variablen zur Treatment-applikation			
KBV	.49	34	.001
DHV	.24	22	.006
KBW	.18	200	.003
PROZRG/8	-.36	39	.005
MIN	-.28	76	.001
ZUG	-.33	155	.001
INST	.36	97	.001
(D) Stichprobenmerkmale			
Stichprobengröße	.24	242	.001
Alter	-.48	245	.001
Geschlecht	-.16	255	.001
SES	-.32	65	.004

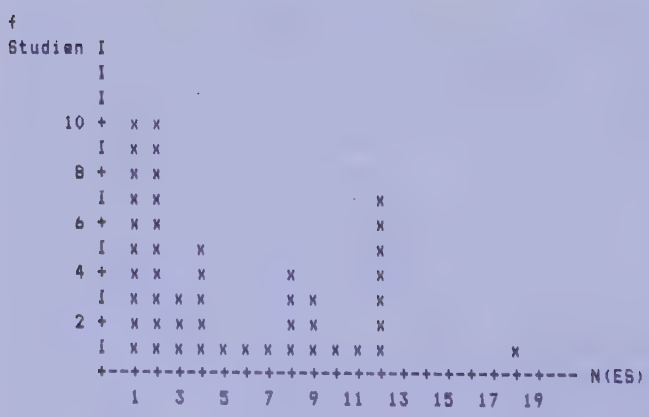
UV = unabhängige Variable; AV = abhängige Variable; KBV = kontingente Bekräftigung vor dem Meßvorgang; DHV = Diskriminative Hinweise vor dem Meßvorgang; PROZRG/S = Prozentwert kooperativer Reaktionen des Simultanspielers; MIN = Treatmentdauer in Minuten; ZUG = Treatmentumfang in Anzahl der Spielzüge; INST = Instruktion zum Spiel; SES = Sozioökonomischer Status

Dissertationen sind (aufgrund ihres exemplarischen Charakters) wesentlich informativer. Ihre interne Gültigkeit ist höher, ihre externe jedoch niedriger. Wird in Dissertationen der Einfluß von Kontingenzen untersucht, werden solche Bekräftigungsstrukturen implementiert, die eine hohe Rangposition aufweisen. Die Instruktionen zur (kooperativen) Interaktion weist umfangreicher auf Kooperationsmöglichkeiten hin, als dies in Zeitschriftenveröffentlichungen der Fall ist. Ebenfalls werden in Dissertationen andere Populationen untersucht. Die Probanden sind jünger und gehören im Durchschnitt einer niedrigeren sozialen Schicht an als Untersuchungsgruppen in Zeitschriftenbeiträgen.

ANZAHL DER EFFEKTGRÖSSEN PRO STUDIE

In der Regel konnten pro Untersuchung mehrere Effekte kalkuliert werden.

Abb. 14 Häufigkeiten der Studien in Abhängigkeit der Anzahl ermittelter Effekte



Zwischen der Anzahl der Effekte pro Studie und der Effektgroße besteht kein linearer Zusammenhang. Die mittleren Effekte in Abhängigkeit der Anzahl der Effekte pro Studie zeigt Tabelle 38.

Tabelle 40 Effektgröße in Abhängigkeiten der Anzahl der Effekte pro Studie

Effekte pro Studie	AM(ES)	SD	SE	MD	N(ES)
1	1.16	.61	.19	1.03	10
2	1.02	1.05	.24	.91	20
3	.69	.48	.16	.81	9
4	.39	1.17	.26	.1	20
5	.3	.18	.08	.2	5
6	.61	.15	.06	.55	6
7					
8	.88	.88	.16	.68	32
9	.62	.75	.18	.53	27
10	.88	.33	.11	.78	10
11					
12	.44	.64	.07	.34	84
18	.46	.42	.08	.53	18

Die Tabellen 39 bis 43 zeigen die Effektgrößen in Abhängigkeit der Fragestellungen, unter denen die Studien durchgeführt wurden.

Tabelle 39: Differenzierung der Kategorie KONTINGENTE BEKRAFTIGUNG

Inhalt der Fragestellung	Studie	AM(ES)	SD	N(ES)	ES max.	ES min.
Welche Kontingenzen wirken?	0101	1.02	-	1	-	-
	0801	.88	.51	8	1.49	.2
	1202	.52	.2	9	.75	.08
	1301	1.71	.3	8	1.93	1.01
	1304	1.57	.54	2	1.95	1.19
	1401	3.21	.34	2	3.45	2.97
	1801	1.23	-	1	-	-
	2201	.65	-	1	-	-
Wirken soziale Verstärker?	0202	1.56	.18	2	1.69	1.43
	1101	1.11	-	1	-	-
	1901	.65	1.51	7	3.75	- .71
	1902	.93	.39	3	1.33	.56
Erhöht die Kenntnis der Kontingenz Kooperation?	2202	1.65	-	1	-	-
Beeinflußt der Umtausch-wert die Kooperation?	0701	.3	.18	5	.57	.15
	1310	.88	.33	10	1.4	.36
Welche Wirkung zeigt differentielle Bekräftigung?	0203	.99	-	1	-	-

Tabelle 40: Differenzierung der Kategorie MODELLERNEN

Inhalt der Fragestellung	Studie	AM(ES)	SD	N(ES)	ES max.	ES min.
Wirkt Modellernen bei der Ausformung von Kooperation?	0804	.01	.21	2	.16	- .14
	0901	.64	.85	2	1.24	.05
	2002	2.58	-	1	-	-
	2003	.18	.92	12	1.64	-1.27
Modellierung der Kontingenz?	0201	.32	.52	9	1.16	- .39
	0601	1.13	2.44	4	4.03	-1.53
Modellierung der Reaktion	0206	- .30	.41	4	.17	- .76
	0301	1.3	-	1	-	-
	1601	.51	.68	3	.99	- .27
Modellierung der Spielstrategie (PD8)	0205	.49	.07	2	.49	.48
	1203	.61	.29	8	1.14	.22

Tabelle 41: Differenzierung der Kategorie VARIATION DER MATRX-WERTE

Inhalt der Fragestellung	Studie	AM(ES)	SD	N(ES)	ES max.	ES min.
Führt die Variation der Matrixwerte zur Variation der kooperativen Wahlen	0702	.63	.41	3	.92	.17
	1307	.31	.34	12	.75	- .30
	1501	1.39	.47	11	2.09	.31
	2301	1.39	.38	12	2.25	.98

Tabelle 42: Differenzierung der Variable HINWEISE ZUM KO-AKTEUR

Inhalt der Fragestellung	Studie	AM(ES)	SD	N(ES)	ES max.	ES min.
Beeinflußt die Informa-	0703	.77	.17	12	1.06	.53
tion über das Verhalten	1001	.54	-	1	-	-
des Ko-Akteurs	1201	.61	.15	6	.92	.52
	1903	.21	.77	2	.75	- .34
	1904	.19	.31	12	.95	- .22
Gruppennormen	1501	1.39	.47	11	2.09	.31
Spiel-Instruktion	2204	1.05	.21	2	1.19	.90
Kooperationsrate des Simultanspielers	1305	.16	.7	4	1.19	- .37
Vertrauensinduktion	1502	.31	.34	12	.99	- .17

Tabelle 43: Differenzierung der Variablen AUFGABENSTRUKTUR

Inhalt der Fragestellung	Studie	AM(ES)	SD	N(ES)	ES max.	ES min.
Transparenz der Aufgabe	2001	-.07	1.09	2	.7	-.85
Aufgabenniveau	0204	.18	.07	4	.23	.09
Aufgabenstruktur	1204	.08	.31	12	.43	-.42

Wenn auch die Fragestellung bereits auf die entsprechende Referenztheorie häufig hinweist, zeigt Abbildung 15, daß nicht unbedingt vom Inhalt der Fragestellung auf die Theorie, mit deren Hilfe die Frage beantwortet werden soll, geschlossen werden kann.

Abb. 15: Beziehung zwischen der Kategorie der Fragestellung und der Kategorie der Referenztheorie am Beispiel 'kontingenter Bekräftigung'

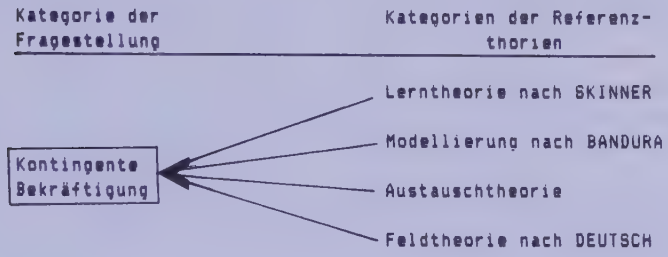


Tabelle 44 weist die mittleren Effektgrößen pro Studie in Abhängigkeit der Referenztheorien aus.

Tabelle 44: Effektgrößen in Abhängigkeit der Referenztheorien

Referenztheorie	Studie	AM(ES)	SD	N(ES)	ES max.	ES min.
OP (SKINNER)	0101	1.02	-	1	-	-
	0202	1.56	.18	2	1.69	1.43
	0203	.99	-	1	-	-
	0801	.88	.51	8	1.49	.2
	0803	.79	.57	4	1.25	- .03
	1101	1.11	-	1	-	-
	1202	.52	.2	9	.75	.08
	1304	1.57	.54	2	1.95	1.19
	1308	.37	.44	18	1.48	- .24
	1401	3.21	.34	2	3.45	2.97
	1801	1.23	-	1	-	-
	1901	.65	1.51	7	3.75	- .71
	1902	.93	.39	3	1.33	.56
	2201	.65	-	1	-	-
	2202	1.65	-	1	-	-
	2203	.57	-	1	-	-
	2204	1.05	.21	2	1.19	.9
Zusammenfassung	18 Stud.	.79	.52	74	3.75	- .71

Fortsetzung Tab. 44

Referenztheorie	Studie	AM(ES)	SD	N(ES)	ES max.	ES min.
SKLT (BANDURA)	0201	.32	.52	9	1.16	- .39
	0205	.49	-	2	.49	.48
	0206	- .3	.41	4	.17	- .76
	0301	1.3	-	1	-	-
	0601	1.13	2.44	4	4.03	-1.53
	0901	.64	.85	2	1.24	.05
	1601	.51	.68	3	.99	- .27
	2002	2.58	-	1	-	-
Zusammenfassung	8 Stud.	.53	.61	26	4.03	-1.53
AT	1305	.16	.70	4	1.19	- .37
	1306	.52	.37	8	1.07	.0
	1501	1.39	.47	11	2.09	.31
Zusammenfassung	3 Stud.	.87	.52	23	2.09	- .37
FTD	0204	.18	.07	4	.23	.09
	0701	.3	.18	5	.57	.15
	0702	.63	.41	3	.92	.17
	0703	.77	.17	12	1.06	.53
	1203	.61	.29	8	1.14	.22
	1204	.08	.31	12	.43	- .42
	1301	1.71	.3	8	1.93	1.01
	1307	.31	.34	12	.75	- .3
	1310	.88	.33	10	1.4	.36
	1502	.31	.34	12	.99	- .17
	1903	.21	.77	2	.75	- .34
	1904	.19	.31	12	.95	- .22
	2001	- .07	1.09	2	.7	- .85
	2301	1.39	.38	12	2.25	.98
Zusammenfassung	14 Stud.	.6	.5	114	2.25	- .85
Rest	1001	.54	-	1	-	-
	1309	1.56	.22	2	1.71	1.4

Während Untersuchungen in Referenz der Austauschtheorie(n) die höchsten Effekte liefern, schneiden Untersuchungen in Referenz der sozial-kognitiven Theorie BANDURAS am schlechtesten ab. Auffallend ist weiter die unterschiedliche Anzahl kalkulierbarer Effekte pro Studie. Austauschtheoretisch und feldtheoretisch orientierte Untersuchungen ermöglichen die Extraktion nahezu doppelt so vieler Effektgrößen wie lerntheoretisch untersuchte Studien. Dies ist insbesondere auf die Variation der Matrixindizes zurückzuführen.

Innerhalb jeder Referenztheorie wurde zwischen verschiedenen Hypothesen und ihnen zuordbaren Konstrukten unterschieden. Die nachfolgende Übersicht zeigt die Effekte in Abhängigkeit dieser Konstrukte.

Tabelle 45 : Effektgrößen in Abhängigkeit der Konstrukte verschiedener Theorien

Referenztheorie	Konstrukte	AM(ES)
Operante Konditionierung	Kontingente Bekräftigung kooperativen Verhaltens	.84
	Kontingente Bekräftigung alternativer Reaktionen	1.01
	Wahrnehmung der Bekräftigungskontingenz	1.25
	Modellierung	.37
Modellierung	der Kontingenz	.66
	kooperativer Reaktionen	.19
Austauschtheorie	Kontingente Bekräftigung kooperativen Verhaltens	.52
	Kontingente Bekräftigung alternativer Reaktionen	1.06
Feldtheorie	Kontingente Bekräftigung kooperativer Reaktionen	.69
	Vertrauen	.75
	Motivinduktion	.43

EFFEKTGRÖSSEN IN ABHÄNGIGKEIT DER KODIERTEN UNABHANGIGEN VARIABLEN

Während die bisherigen gemittelten Effektgrößen in Abhängigkeit des Konstrukts gebildet wurden, liefert die nachfolgende Tabelle die mittleren Effekte in Abhängigkeit der a posteriori transformierten Variablen auf der operationalen Ebene.

Tabelle 46: Effektgrößen in Abhängigkeit der a posteriori transformierten Variablen

Prädiktorvariable	N(ES)	AM(ES)	SD	SE	MD
Modellernen	42	.31	.99	.15	.16
Bekräftigungskontingenzen für Kooperation	34	1.12	.82	.14	.82
Bekräftigung alternativer Reaktionen	22	.76	.82	.18	.71
Diskriminative Hinweise	89	.84	.83	.09	.77
Bekräftigungskontingenzen während des Meßvorgangs	200	.76	.68	.05	.67
Tit for Tat	23	.86	.72	.15	1.16
PROZRG/S	39	.55	.49	.08	.46
Display	22	.41	.40	.09	.51
Diskriminative Hinweise während des Meßvorgangs	185	.66	.69	.05	.59
Revenuen	195	.70	.67	.05	.61

Die Mittelwerte informieren nicht über die Wirkung der einzelnen Variablen, sondern nur darüber, daß unabhängig davon, welche Variablen noch beteiligt sind, nach diesen Prädiktoren geblockt, die entsprechenden Effekte entstehen.

Die Wirkung der einzelnen Niveaus des Modellernens zeigt Abb 16.

Abb. 16: Effektgrößen für die Kategorie Modellernen unter Berücksichtigung konfundierter Variablen

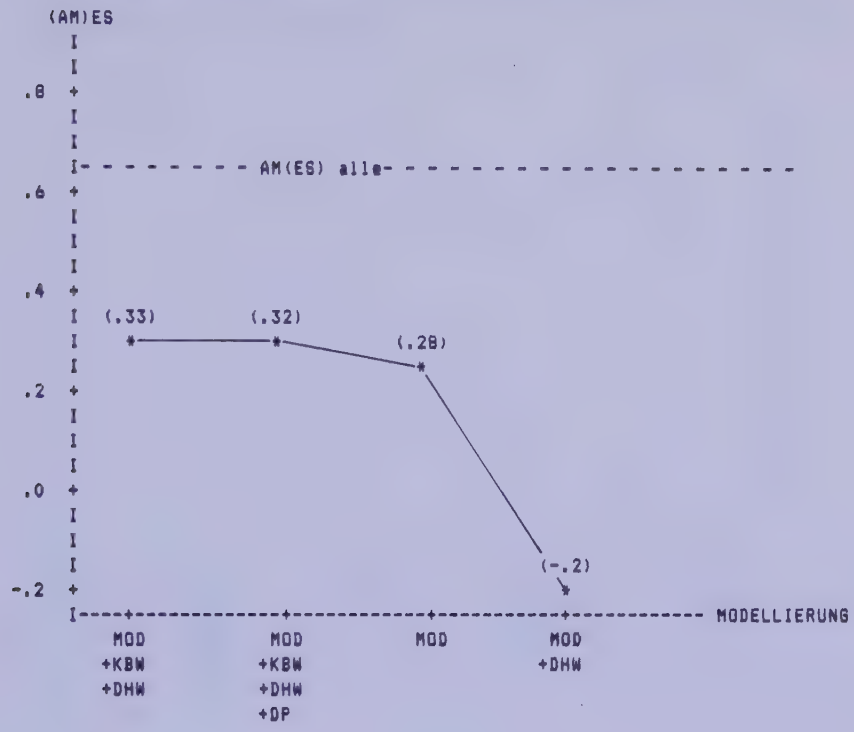
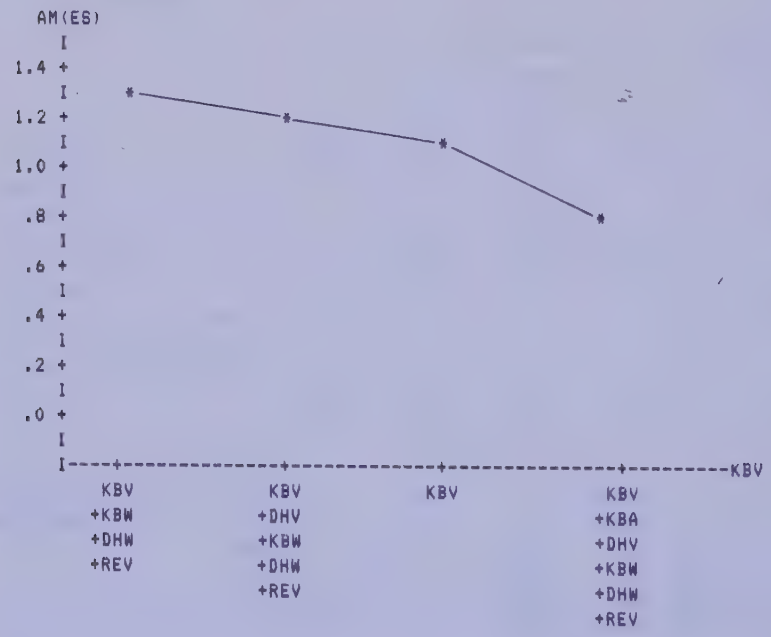


Abb. 16 zeigt die Wirkung des Modellernens in Abhängigkeit verschiedener applizierter oder moderierender Variablen. Die Modellierung kooperativer Reaktionen zeigt dann Wirkung, wenn während des Meßvorgangs die kooperative Kontingenz belohnt wird. In Analogie zur Literaturdiskussion führt die Anzeige des Gewinns des Ko-Akteurs im Vergleich zum eigenen Gewinn zur Abnahme des Effekts. Für Modell-Situationen sind diskriminative Hinweise auf interferente Situationen und niedrige Revenuen in Form von Punkten die ungünstigsten Voraussetzungen zur Ausformung kooperativen Verhaltens.

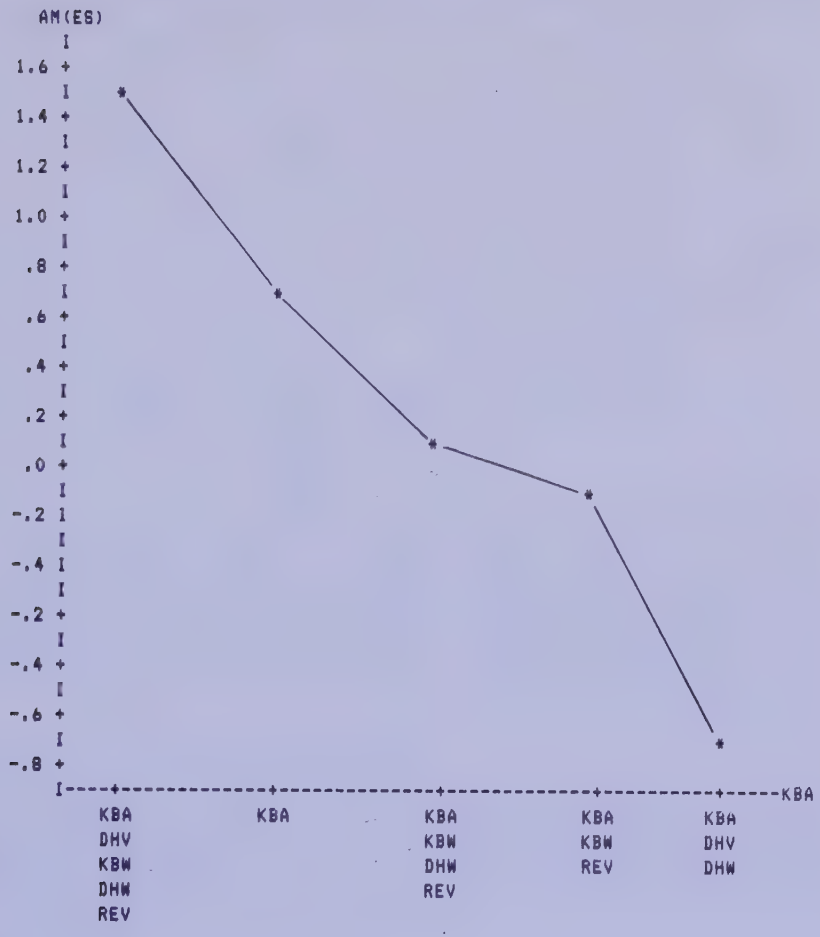
Tabelle 44 weist bereits darauf hin, daß Experimente, die lentheoretisch orientierte Operationen anwenden, zu hohen Effektgrößen führen.

Abb. 17: Mittlere Effektgrößen für KBV unter Berücksichtigung weiterer Variablen



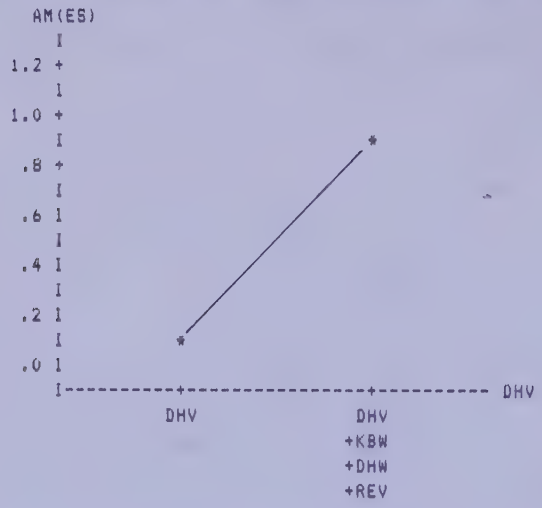
Solche Treatments sind besonders wirksam, die neben der kontingenten Bekräftigung die Kooperativität der Situation anzeigen und die Adäquanz des Verhaltens während des Meßvorgangs verstärkt wird. Werden zusätzlich alternative Reaktionen bekräftigt, sinkt die Effektgröße. Abbildung 18 informiert detailliert über den Einfluß kontingenter Bekräftigung alternativer Reaktionen. Multiple, alternative Reaktionen bekräftigende Kontingenzen führen zu einem Effekt, der nur dann 'überboten' wird, wenn sowohl Information über die Kooperation der Situation angeboten wird und während des Meßvorgangs umfangreiche, die Kooperation unterstützende Variablen wirken. Den diskriminativen Hinweisen zur Kooperativität der Situation kommt dabei vor dem Meßvorgang eine weniger wichtige Funktion zu: sie beeinflussen während des Meßvorgangs die Situation. Werden alternative Reaktionen höher bekräftigt als kooperative Reaktionen, sinkt das Kooperationsniveau.

Abb. 18: Mittlere Effekte für KBA in Abhängigkeit weiterer Variablen



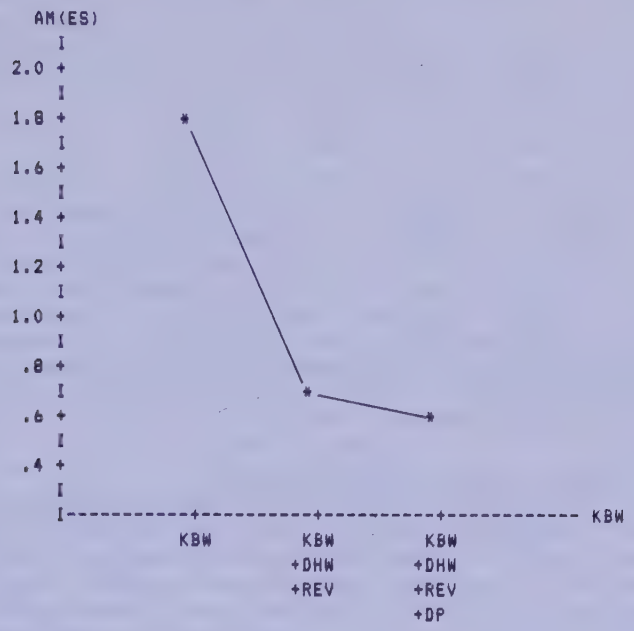
Diskriminative Hinweise als alleinige Treatmentbedingung vor dem Meßvorgang führt selbst dann, wenn die Hinweise eindeutig die Kooperativität der Situation anzeigen, zu niedrigen Effekten (AM(ES) = .12 sigma).

Abb. 19: Mittlere Effektgrößen für DHV in Abhängigkeit weiterer Variablen



Erst Hinweise während des Meßvorgangs erhöhen erheblich die Treatmentwirkung selbst dann, wenn alternative Raktionen (niedrig) belohnt werden und die Revenuen niedrig ausfallen. Die Kombination diskriminativer Hinweise und Bekräftigungskontingenzen ausschließlich kooperativer Reaktionen während des Meßvorgangs führen zu hohen Effekten. Diesen Zusammenhang weist die folgende Abbildung aus.

Abb 20: Mittlere Effekte von KBW in Abhängigkeit weiterer Variablen



FEINANALYSE DER TRANSFORMIERTEN VARIABLEN

Die bisherigen Ergebnisse bilden die Wirkung einzelner und konfundierter transformierter Variablen ab, ohne die jeweilige Variable näher zu differenzieren. Die folgende Feinanalyse zeigt demgegenüber mittlere Effektgrößen in Abhängigkeit der durch die Rangpositionen unterschiedlich gewichteten Niveaus innerhalb der Variablen.

MODELLERNEN

Experimente, die Kooperation durch Modellernen zu erhöhen beabsichtigen, zeigen niedrige Effekte. Probanden, die Modelle betrachten, die eine TFT-Strategie ohne nähere Erläuterung spielen, zeigen sogar eine niedrigere Kooperationsrate als die Kontrollgruppe oder aber fallen im Verlauf der Spielzüge in ihrer Kooperationsrate ab. Modelle, die insgesamt mehr nichtkooperative Reaktionen äußern, führen zu keinem Effekt. Erst die Darstel-

lung kooperativer Interaktionen in Form kooperativer Episoden führt zu einem mittleren Kooperationseffekt. Wird zusätzlich die Kontingenz, unter der die kooperative Interaktion steht, den Beobachtern vermittelt, steigt die Wirkung an. Jedoch führt erst die Modellierung individuell bekräftigter kooperativer Episoden bei gleichzeitiger Bestrafung alternativer Reaktionen zu einem wahrnehmbaren höheren Erfolg in der Akquisitionsphase. Tabelle 47 weist auf die Rangunterschiede zwischen der theoretisch festgelegten Rangfolge und den empirisch gefundenen Rangpositionen.

KONTINGENTE BEKRÄFTIGUNG KOOPERATIVER REAKTIONEN VOR DEM MESSVORGANG
Der mittlere Effekt für Untersuchungen, die kontingente Bekräftigung kooperativen Verhaltens als Treatment in vivo applizieren, zeigt einen wesentlichen Unterschied zum über-alles-Effekt. Tabelle 48 zeigt leichte Unterschiede bzw. eine positive Korrelation zwischen Intensität der Bekräftigung und den Effektgrößen. Entgegen der Vermutung liefert die Bekräftigung prosozialer Verhaltensweisen, unter denen Kooperation eine Kategorie darstellt, höhere Effekte als die individuelle Bekräftigung kooperativer Episoden. Gruppenkontingente Bekräftigung koopertiver Episoden führt - in Analogie zur Diskussion in der Literatur - zu den höchsten und zu homogenen Effekten.

KONTINGENTE BEKRÄFTIGUNG ALTERNATIVER REAKTIONEN
Innerhalb der transformierten Variablen erzeugt diese Kategorie enorme Unterschiede, wenngleich der mittlere Effekt nicht wesentlich vom über-alles-Effekt abweicht (vgl. Tabelle 49). Erwartungsgemäß sinkt der mittlere Effekt unter ES = 0 sigma, wenn alternative Reaktionen höher belohnt werden als kooperatives Verhalten. Alleine die Möglichkeit, kooperatives Verhalten zu bestrafen, reduziert nicht die Kooperationsbereitschaft wesentlich. Entgegen der Erwartung steigt die Kooperationsbereitschaft sogar an, wenn die Möglichkeit besteht, kooperative Reaktionen hoch zu bestrafen.

DISKRIMINATIVE HINWEISE
In der Literatur wird der Einfluß diskriminativer Hinweise zur (Un-)Kooperativität der Person und/oder der Situation extensiv diskutiert. Werden solche Hinweise vor dem Meßvorgang vorgegeben und bleiben alle anderen Einflüsse unberücksichtigt, liegt der mittlere Effekt über dem über-alles-Effekt (siehe Tab.50). Dieses Ergebnis überrascht, da die Variable hier auch Informationen zur Un-Kooperativität der Person und/oder Situation erfaßt. Es ist zu vermuten, daß allein die Information zur Situation diskriminative Hinweise liefert, die insgesamt kooperationsfördernd wirken. Mehrdeutige oder unzureichende Hinweise drücken jedoch die Höhe der Kooperationsbereitschaft.

KONTINGENZEN WÄHREND DES MESSVORGANGS
Diese Variable wurde aufgrund der Analyse der operationalen Vorgehensweise kodiert, unabhängig davon, ob die Autoren explizit Bekräftigungskontingenzen während des Meßvorgangs zu kontrollieren beabsichtigten. Dies ist insofern konform mit theoretischen Annahmen, die externe Kontingenzen manipu-

lieren, ohne die Kontingenzen aufgrund der sozialen Interaktion zu kontrollieren. Tabelle 51 zeigt dann auch einen niedrigen Effekt für solche Experimente, in denen keine Bekräftigungskontingenzen kooperativen Verhaltens vorliegen (AM(ES) = .28) bzw. in denen kooperatives Verhalten bestraft wird (AM(ES = -.07s). Überraschenderweise fällt der Effekt für Kontingenzen, in denen kooperatives Verhalten niedriger belohnt wird als nicht-kooperatives Verhalten höher aus (AM(ES) = .72s) als für solche Kontingenzen, in denen die Bekräftigung für kooperatives Verhalten niedriger ausfällt als für nicht-kooperatives Verhalten. Während für Untersuchungen, die die Möglichkeit zur Bestrafung kooperativen Verhaltens als Treatment vor dem Meßvorgang eröffnen, hoch ausfällt (AM(ES) = 1.29s), liefern Studien, in denen während des Meßvorgangs Kooperation bestraft wird, einen niedrigen Effekt (AM(ES) = -.01s). Demnach wirkte ein Bestrafungsmechanismus in die umgekehrte Richtung. Unter der Prämisse, daß zur Zielerreichung Kooperation notwendig ist, verhindert die Bestrafung des Ko-Akteurs die Zielerreichung (und vice versa), so daß während des Meßvorgangs solche interdependenten Kontingenzen vorherrschen, die eine hohe Auftrittswahrscheinlichkeit kooperativer Reaktionen zur Folge haben. Eine, in dieselbe Richung zielende Beziehung zeigen solche Treatments, in denen vor dem Meßvorgang die Belohnung für nicht-kooperative Reaktionen größer ausfällt als für kooperative (AM(ES) = -.12)1 während des Meßvorgangs wird das rationale Verhalten, die kooperative Reaktion, auch dann ausgeführt, wenn die (hohe) Belohnung alternativer Reaktionen möglich ist. Die Effektgrößen liegen über ES gt 1s, wenn die kooperative Reaktion vor und während des Meßvorgangs bekräftigt werden.

Abb. 21: Effektgrößen in Abhangigkeit verschiedener Modellierungs-Niveaus

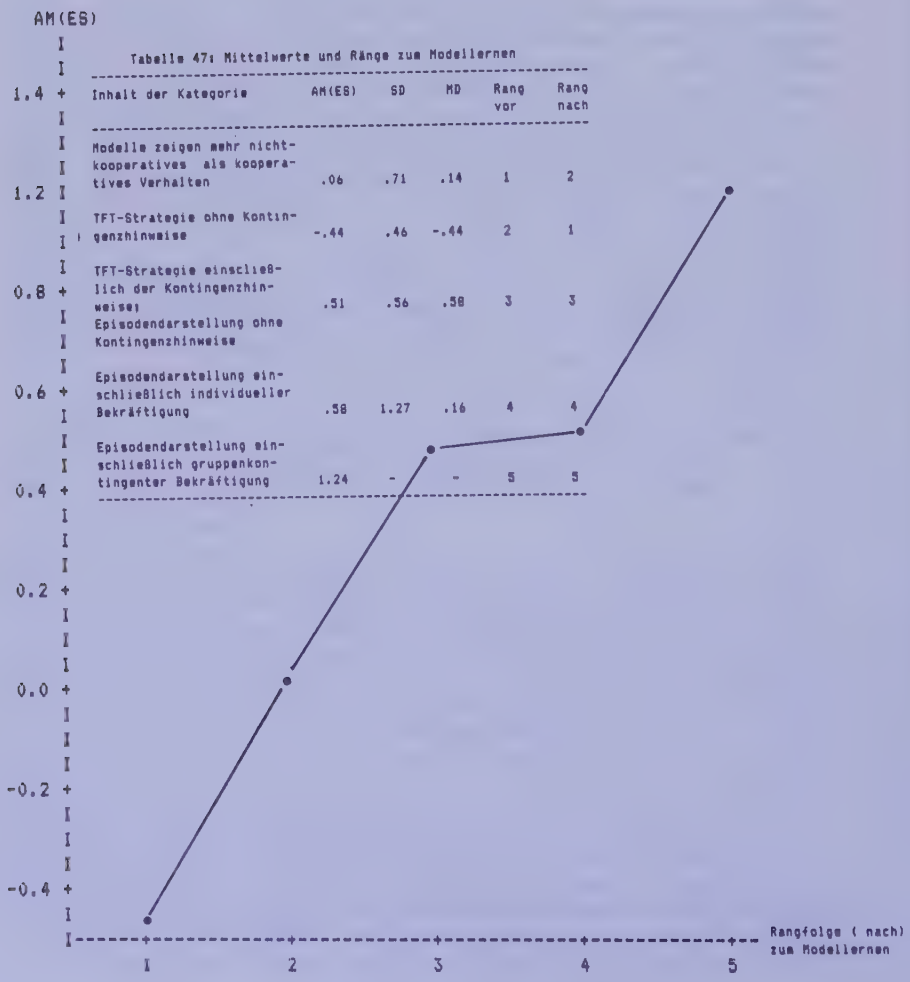


Tabelle 47: Mittelwerte und Ränge zum Modellernen

Inhalt der Kategorie	AM(ES)	SD	MD	Rang vor	Rang nach
Modelle zeigen mehr nichtkooperatives als kooperatives Verhalten	.06	.71	.14	1	2
TFT-Strategie ohne Kontingenzhinweise	-.44	.46	-.44	2	1
TFT-Strategie einschließlich der Kontingenzhinweise) Episodendarstellung ohne Kontingenzhinweise	.51	.56	.58	3	3
Episodendarstellung einschließlich individueller Bekräftigung	.58	1.27	.16	4	4
Episodendarstellung einschließlich gruppenkontingenter Bekräftigung	1.24	-	-	5	5

Abb. 22: Effektgrößen in Abhängigkeit verschiedener Niveaus kontingenter Bekräftigung vor dem Meßvorgang

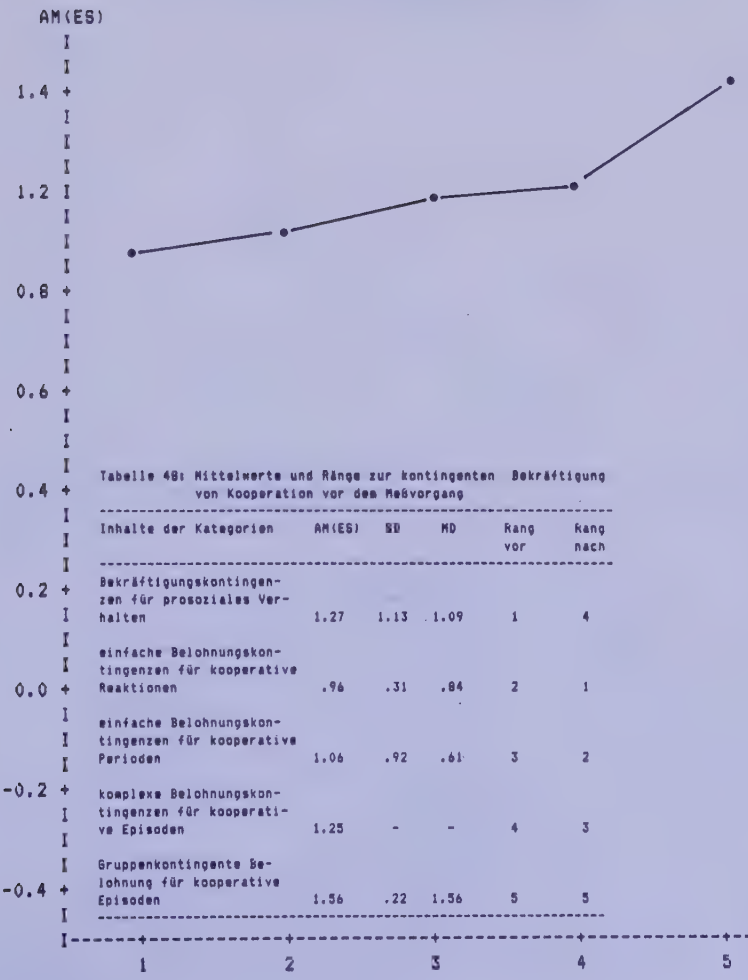


Tabelle 48: Mittelwerte und Ränge zur kontingenten Bekräftigung von Kooperation vor dem Meßvorgang

Inhalte der Kategorien	AM(ES)	SD	MD	Rang vor	Rang nach
Bekräftigungskontingenzen für prosoziales Verhalten	1.27	1.13	1.09	1	4
einfache Belohnungskontingenzen für kooperative Reaktionen	.96	.31	.84	2	1
einfache Belohnungskontingenzen für kooperative Perioden	1.06	.92	.61	3	2
komplexe Belohnungskontingenzen für kooperative Episoden	1.25	-	-	4	3
Gruppenkontingente Belohnung für kooperative Episoden	1.56	.22	1.56	5	5

Abb. 23: Effektgrößen in Abhängigkeit verschiedener Niveaus kontingenter Bekräftigung alternativer Reaktionen vor dem Meßvorgang

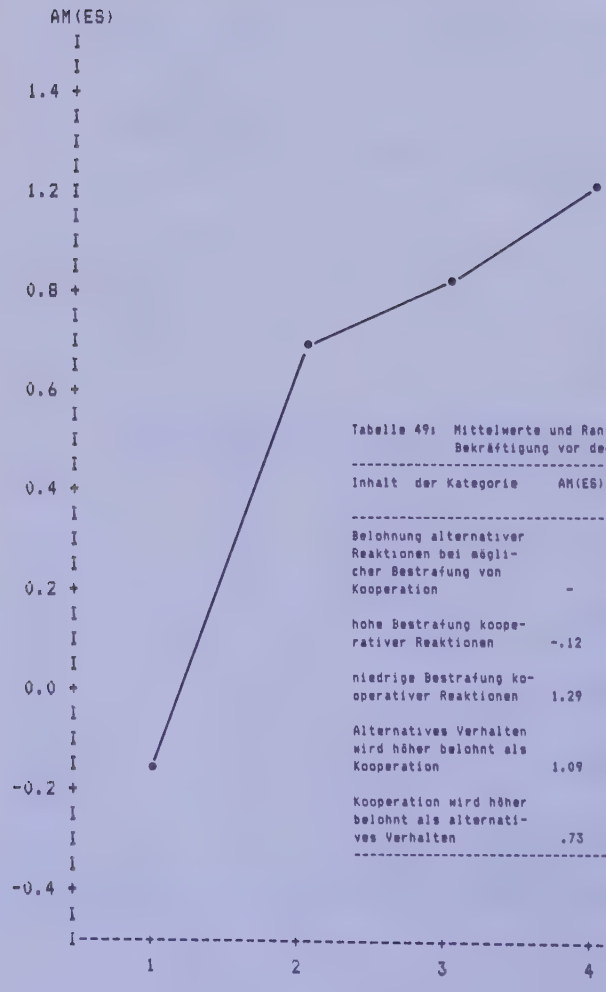


Tabelle 49: Mittelwerte und Rangpositionen zu alternativer Bekräftigung vor dem Meßvorgang

Inhalt der Kategorie	AM(ES)	SD	MD	Rang vor	Rang nach
Belohnung alternativer Reaktionen bei möglicher Bestrafung von Kooperation	-	-	-	1	-
hohe Bestrafung kooperativer Reaktionen	-.12	.55	-.22	2	1
niedrige Bestrafung kooperativer Reaktionen	1.29	.59	1.32	3	4
Alternatives Verhalten wird höher belohnt als Kooperation	1.09	.7	1.2	4	3
Kooperation wird höher belohnt als alternatives Verhalten	.73	.7	.64	5	2

Abb. 24: Effektgrößen in Abhängigkeit verschiedener Niveaus diskriminativer Hinweise vor dem Meßvorgang

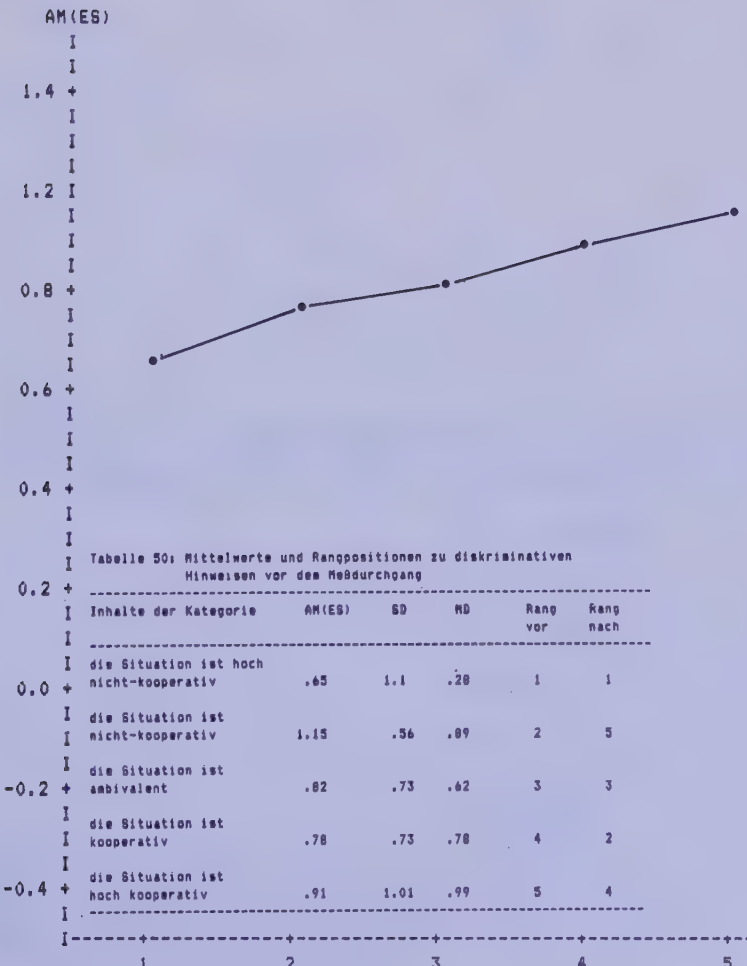


Tabelle 50: Mittelwerte und Rangpositionen zu diskriminativen Hinweisen vor dem Meßdurchgang

Inhalte der Kategorie	AM(ES)	SD	MD	Rang vor	Rang nach
die Situation ist hoch nicht-kooperativ	.65	1.1	.28	1	1
die Situation ist nicht-kooperativ	1.15	.56	.89	2	5
die Situation ist ambivalent	.82	.73	.62	3	3
die Situation ist kooperativ	.78	.73	.78	4	2
die Situation ist hoch kooperativ	.91	1.01	.99	5	4

Abb. 25: Effektgrößen in Abhängigkeit verschiedener Niveaus kontingenter Bekräftigung während des Meßvorgangs

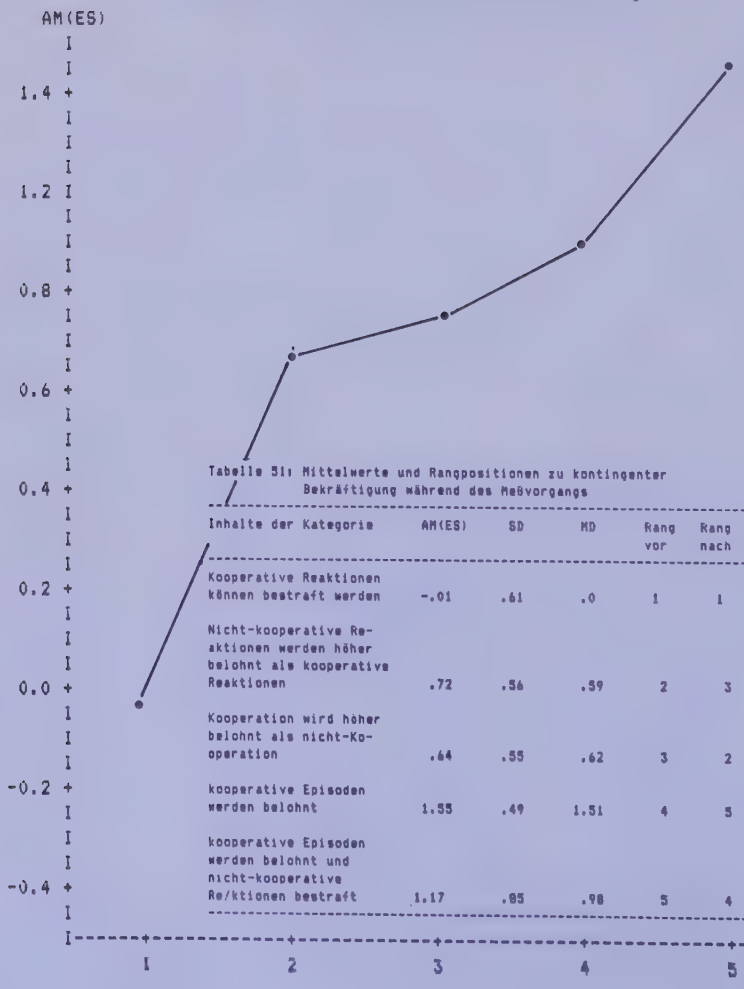
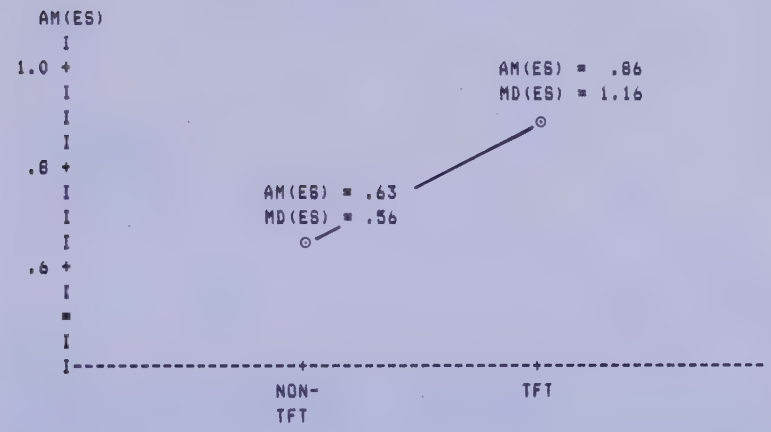


Tabelle 51: Mittelwerte und Rangpositionen zu kontingenter Bekräftigung während des Meßvorgangs

Inhalte der Kategorie	AM(ES)	SD	MD	Rang vor	Rang nach
Kooperative Reaktionen können bestraft werden	-.01	.61	.0	1	1
Nicht-kooperative Reaktionen werden höher belohnt als kooperative Reaktionen	.72	.56	.59	2	3
Kooperation wird höher belohnt als nicht-Kooperation	.64	.55	.62	3	2
kooperative Episoden werden belohnt	1.55	.49	1.51	4	5
kooperative Episoden werden belohnt und nicht-kooperative Reaktionen bestraft	1.17	.65	.98	5	4

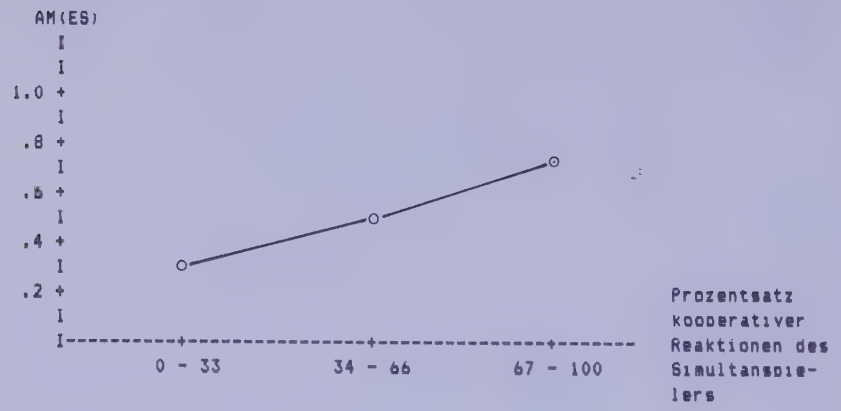
TIT-FOR-TAT (TFT)
Gegenüber der modellierten TFT-Strategie führt die Anwendung dieser Verhaltensstrategie in vivo im PDG zu hohen Effekten (AM(ES) = .86s. MD = 1.16. Sie ist damit im Rahmen der PDG-Forschung die wirksamste Einflußgröße zur Evozierung und Aufrechterhaltung kooperativer Interaktionen.

Abb.26: Effektgrößen in Abhängigkeit von TFT



UNKONTINGENTE RATE KOOPERATIVER REAKTIONEN DURCH DEN SIMULTANEN KO-AKTEUR
Ein strenger Zusammenhang zwischen der unkontingenten Rate kooperativer Verhaltensweisen und Kooperationsverhalten der Probanden besteht nicht.

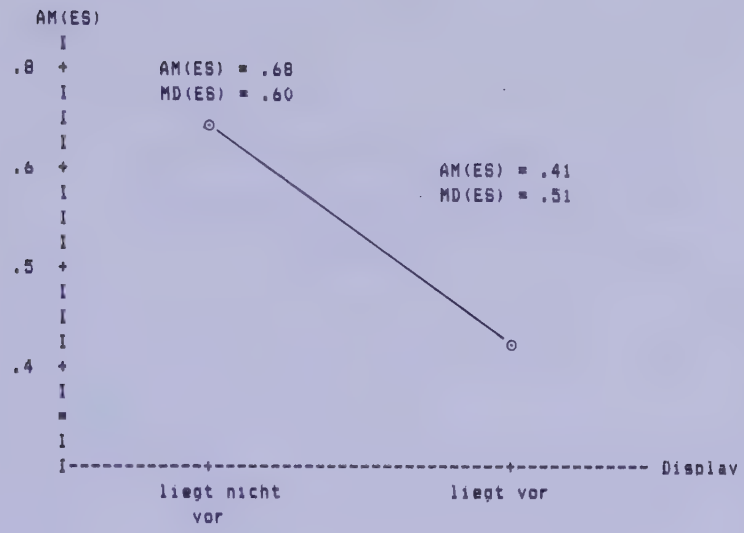
Abb.27: Effektgrößen in Abhangigkeit kooperativer Reaktionen durch den Simultanspieler



ERGEBNISANZEIGE (DISPLAY)

Die kontingente Belohnung kooperativen Verhaltens (in der Form des erreichten Gewinnes, in Form des Verhaltens des Partners oder in Form externer Kontingenzen) kann den Interaktionspartnern angezeigt werden. Differenzen weisen dabei auf eine instabile Reziprozität hin. Temporäre Abweichungen werden ohne die Widerspiegelung der Verhaltenswahlen nicht unbedingt unmittelbar registriert, so daß die Reziprozität mangels differenzierter Wahrnehmung weiter aufrechterhalten bleibt. Die Anzeige der Wirkung des eigenen Verhaltens in Abhängigkeit des Verhaltens des Ko-Akteurs liefert sehr schnell und unmittelbar diskriminative Hinweise, den kooperativen Pfad bei Verletzung der Reziprozität ebenfalls zu verlassen. Abb. 28 zeigt, daß die Monitor-Anzeige der Verhaltensweisen zu einer geringeren Kooperationswahrscheinlichkeit fuhrt.

Abb. 28: Effektgrößen bei Vorlage von Display



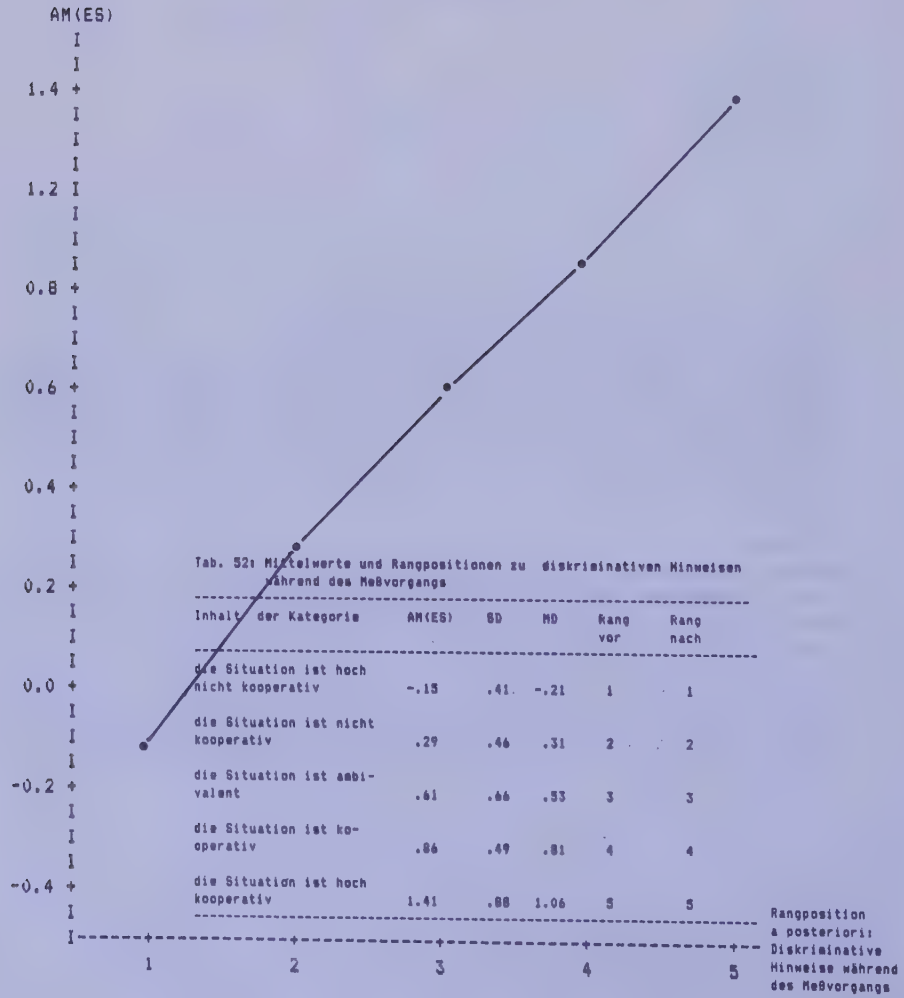
DISKRIMINATIVE HINWEISE WÄHREND DES MESSVORGANGS

Neben der Wiedergabe der Verhaltensinteraktion oder der Ergebnisse dieser Interaktion liefert die kooperative Interdependenz eine Fülle weiterer diskriminativer Hinweise (vgl. weiter vorne Tabelle 42). Abb. 29 zeigt, daß die Hinweise zur Kooperativität die eigene Kooperationsbereitschaft stark beeinflussen.

REVENUEN

Während die Literatur - und infolge dessen die Rangkodierung - einen positiven Zusammenhang zwischen Revenuen-Niveau und Kooperation unterstellen, zeigt Abb. 30 den umgekehrten Zusammenhang. Danach reduzieren hohe Revenuen keine Interferenzen zur Kooperation, sondern scheinen sie eher zu verstärken. Es ist zu vermuten, daß je höher die Differenz zwischen Kosten und Nutzen psychisch interpretiert wird, desto geringer die Wahrscheinlichkeit eingeschätzt wird, eine kooperative Interaktion könne bestehen bzw. aufrechterhalten bleiben - oder auf der Ebene der Episode - desto unwahrscheinlicher reagiert der Partner ebenfalls mit gebenden Reaktionen.

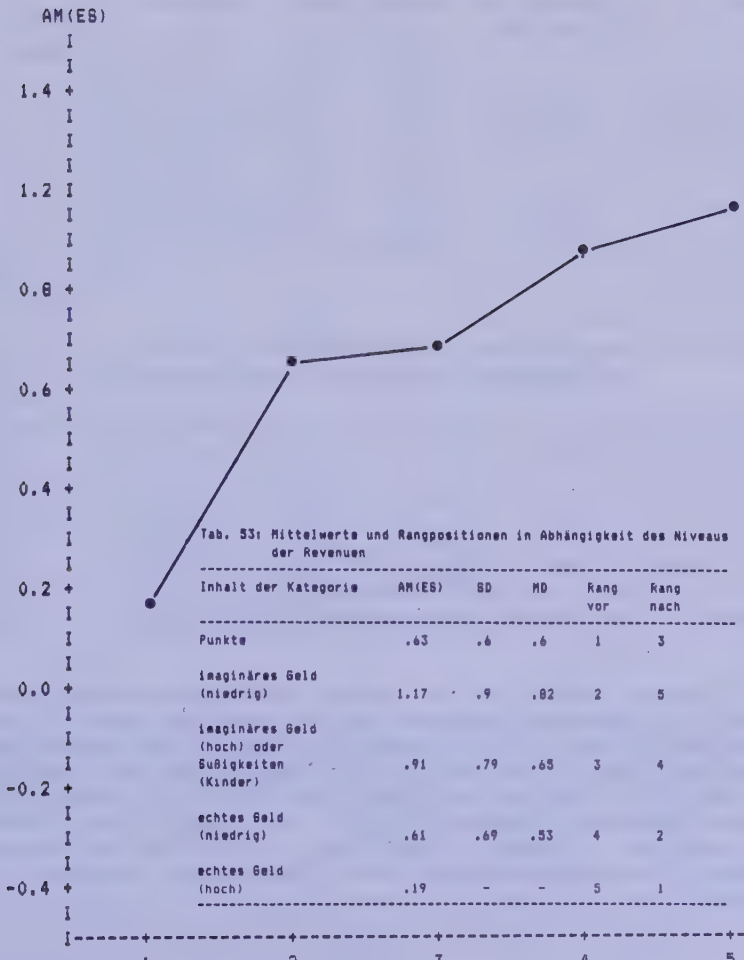
Abb. 29: Effektgrößen in Abhängigkeit verschiedener Niveaus diskriminativer Hinweise während des Meßvorgangs



Tab. 52: Mittelwerte und Rangpositionen zu diskriminativen Hinweisen während des Meßvorgangs

Inhalt der Kategorie	AM(ES)	SD	MD	Rang vor	Rang nach
die Situation ist hoch nicht kooperativ	-.15	.41	-.21	1	1
die Situation ist nicht kooperativ	.29	.46	.31	2	2
die Situation ist ambivalent	.61	.66	.53	3	3
die Situation ist kooperativ	.86	.49	.81	4	4
die Situation ist hoch kooperativ	1.41	.88	1.06	5	5

Abb. 30: Effektgrößen in Abhängigkeit verschiedener Revenuen



Tab. 53: Mittelwerte und Rangpositionen in Abhängigkeit des Niveaus der Revenuen

Inhalt der Kategorie	AM(ES)	SD	MD	Rang vor	Rang nach
Punkte	.63	.6	.6	1	3
imaginäres Geld (niedrig)	1.17	.9	.82	2	5
imaginäres Geld (hoch) oder Süßigkeiten (Kinder)	.91	.79	.65	3	4
echtes Geld (niedrig)	.61	.69	.53	4	2
echtes Geld (hoch)	.19	-	-	5	1

MATRIX- UND ZELLENINDIZES

Der Einfluß der Zellenwerte im PDG wurde bereits in seiner Diskussion dargestellt. Die in die Meta-Analyse einfließenden Studien im Rahmen des PDG sind - soweit sie die Zellenindices variieren - in solche zu unterscheiden, die allein die Zellenindicies variieren (Naturspiel) und in solche, die darüber hinaus zusätzliche Treatmentbedingungen applizieren. Die Tabellen 54 und 55 zeigen den Zusammenhang zwischen Zellenwerten und Effektgrößen im Naturspiel und in allen Untersuchungen zum PDG.

Tabelle 54: Spearman-Korrelationen zwischen Zellenwerten und Effektgrößen

Matrixzelle	AM(MI)*	SD(MI)	r(s)	n	a
Re	5.2	3.5	-.27	124	.001
T	6.4	4.2	-.24	124	.004
S	-2.1	4.3	-.04	124	ns
P	- .5	3.6	-.17	124	.029

Tabelle 55: Spearman-Korrelationen zwischen Zellenwerten und Prozentwert kooperativer Wahlen im Naturspiel

Matrixzelle	AM(MI)*	SD(MI)	r(s)	n	a
Re	5.7	3.6	-.11	29	ns
T	7.6	4.7	-.04	29	ns
S	-2.1	5.2	-.08	29	ns
P	- .07	3.6	-.55	29	.001

* AM(MI) = Mittelwert der entsprechenden Zellen

Liegen - ungeachtet welcher Qualität - zusätzliche Antezendenzen vor, liefert die intra-interaktionale Belohnung (Re) die größte Einflußquelle. Dies ist insofern plausibel, als die Treatments in ihrer Mehrheit kooperatives Verhalten auszuformen versuchen und damit die Probanden auf eine kooperative Interaktion orientieren. Wenn Kooperation erwünscht ist, dann spielt die relative Belohnung eine wesentliche Rolle. Im Gegensatz dazu stehen Kooperationspartner im Naturspiel primär unter dem Einfluß der Bestrafungskontingenz. Je niedriger diese ausfällt, desto größer ist die Kooperationsbereitschaft.

Nun sind die Zellenwerte absolut betrachtet von weniger großem Interesse. Ihre wechselseitigen Einflüsse versuchen verschiedene Zellenindices festzustellen. Tabelle 56 bildet diese Ergebnisse ab.

Tabelle 56: Spearman-Korrelationen zwischen Zellenindices und PROZRG für das Nahrungsspiel und für alle Daten

	IN1	WETT	GRAUS	R1	R2	IN6	ANGST	R4	AAC	APT	ICA	PROZRG	
IN1	-	-.2	-.16	.19	.04	-.23	.47	.44	-.39	.81	-.40	-.32	N
WETT	-.72	-	.49	-.33	.10	.55	-.47	-.5	.86	.14	.99	-.02	A
GRAUS	-.36	.39	-	-.88	-.44	.96	-.33	-.06	.72	.20	.40	-.33	T
R1	.27	-.23	.71	-	.57	-.83	.22	-.05	-.64	-.03	-.31	.48	U
R2	.23	-.17	.56	.85	-	.3	.31	.01	.43	-.05	.44	.15	R
IN6	-.45	.44	.79	-.66	-.51	-	-.26	-.19	.23	.17	.02	-.29	S
ANGST	-.64	-.69	.52	-.28	.27	-.45	-	.6	-.24	.19	-.25	-.22	P
R4	.7	-.71	.5	-.26	.19	-.46	.83	-	-.23	.26	-.44	-.35	I
AAC	-.34	.56	.68	-.37	-.21	.55	-.29	-.35	-	.23	.86	-.16	E
APT	.76	-.15	.04	.01	.02	-.08	.43	.42	-.10	-	.13	-.40	L
ICA	.34	.25	.30	-.04	.00	.18	.05	.13	.33	.56	-	.03	
PRG	-.17	.08	.05	-.05	-.1	.06	-.21	-.08	-.02	-.14	.06	-	

alle Daten

IN1 = X1 + X2
WETT = X3 - X2
GRAUS = lg(10)*((X3 - X2) / (X1 - X4))
R1 = (X1 - X4) / (X3 - X2)
R2 = (X1 - X2) / (X3 - X2)
IN6 = lg(10)*((X1 - X2) / (X1 - X4))
ANGST = lg(10)*((X4 - X2) / (X3 - X2))
R4 = lg(10)*((X3 - X1) / (X3 - X2))
AAC = ((X3 - X4) /2) - ((X1 - X2) / 2))
APT = (X1 + X2 + X3 + X4) / 4
PRG = Prozentwert kooperativer Wahlen

TREATMENTDAUER

Für Handlungsmeßverfahren wurde die Treatmentdauer in Minuten, für PDG-Experimente in Anzahl der Spielzüge berechnet. Tabelle 57 zeigt die durchschnittliche Applikationslänge und die Korrelation zwischen Dauer und Effektgröße.

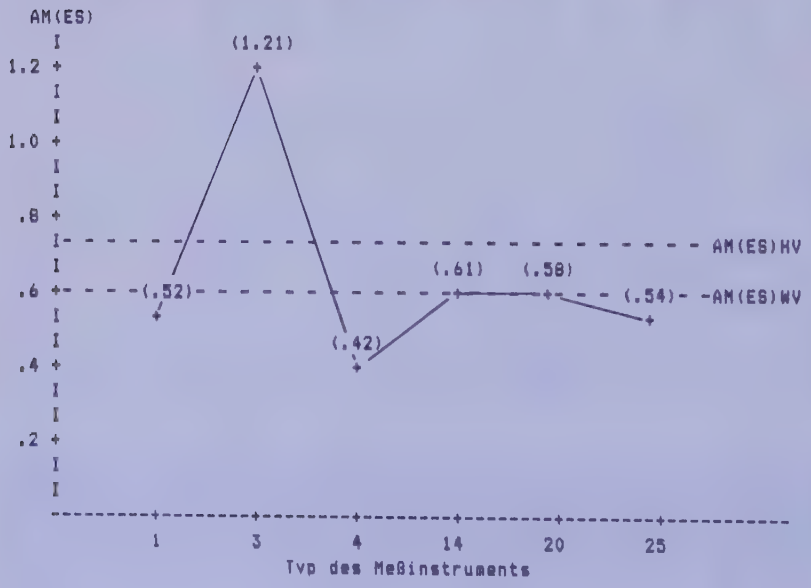
Tabelle 57: Durchschnittlicher Treatmentumfang und Zusammenhang zwischen Treamentdauer und Effektgröße

Dauer in	AM(Dauer)	SD	Dauer max.	Dauer min.	r(s)	N ES
Minuten	58.3	68	300	2	.19	96
Spielzüge	59	70	270	4	.31	148

MESSINSTRUMENT

Die mittleren Effekte variieren in Abhängigkeit des Meßinstrumentes. Studien, die Kooperation über Handlungsinstrumente messen, erzeugen im Mittel eine Effektgröße von AM(ES) = .72 (SD(ES) = 1.03), während Studien im Rahmen der PDG-Forschung einen mittleren Effekt von AM(ES) = .61 (SD(ES) = .53) erzeugen.

Abb. 31: Effektgrößen in Abhängigkeit des Meßinstrumentes



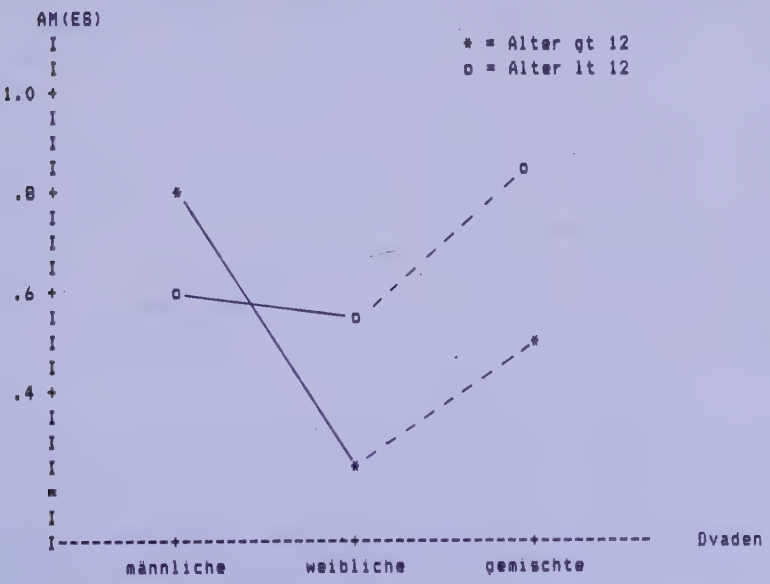
01 = nicht-soziale, interdependente, alternativfreie Teilungsreaktionen
03 = soziale, interdependente, alternativfreie Teilungsreaktionen
04 = soziale, interdependente, alternative Teilungsreaktionen
14 = PDB-Instrumente
20 = Verhaltensbeobachtung
23 = Papier-Bleistifttests
HV = Handlungsverfahren WV = Wahlverfahren

Für 70 Effekte wurde innerhalb der Handlungsverfahren die abhängige Variable auf der Konstruktebene definiert. Für 49 Effekte stimmten Konstrukt und operationale Definition inhaltlich überein.

ALTER UND GESCHLECHT DER PROBANDEN

Zwischen dem Alter der Probanden und der Effektgröße besteht kein linearer Zusammenhang. Kinder im Alter zwischen 5 und 6 Jahren zeigen einen niedrigen Effekt (AM(ES) = .24s), während bei Kindern im Alter zwischen 10 und 12 Jahren ein überdurchschnittlich hoher Effekt (AM(ES) = 1.1s) aufgezeigt werden kann. Zwischen allen Kindern und Erwachsenen besteht kein wesentlicher Unterschied (.67s gegenüber .64s). Frauen zeigen einen niedrigeren Kooperationsanstieg (AM(ES) = .48s) als Männer (AM(ES) = .72s).

Abb. 32: Effektgrößen in Abhängigkeit des Alters und des Geschlechts

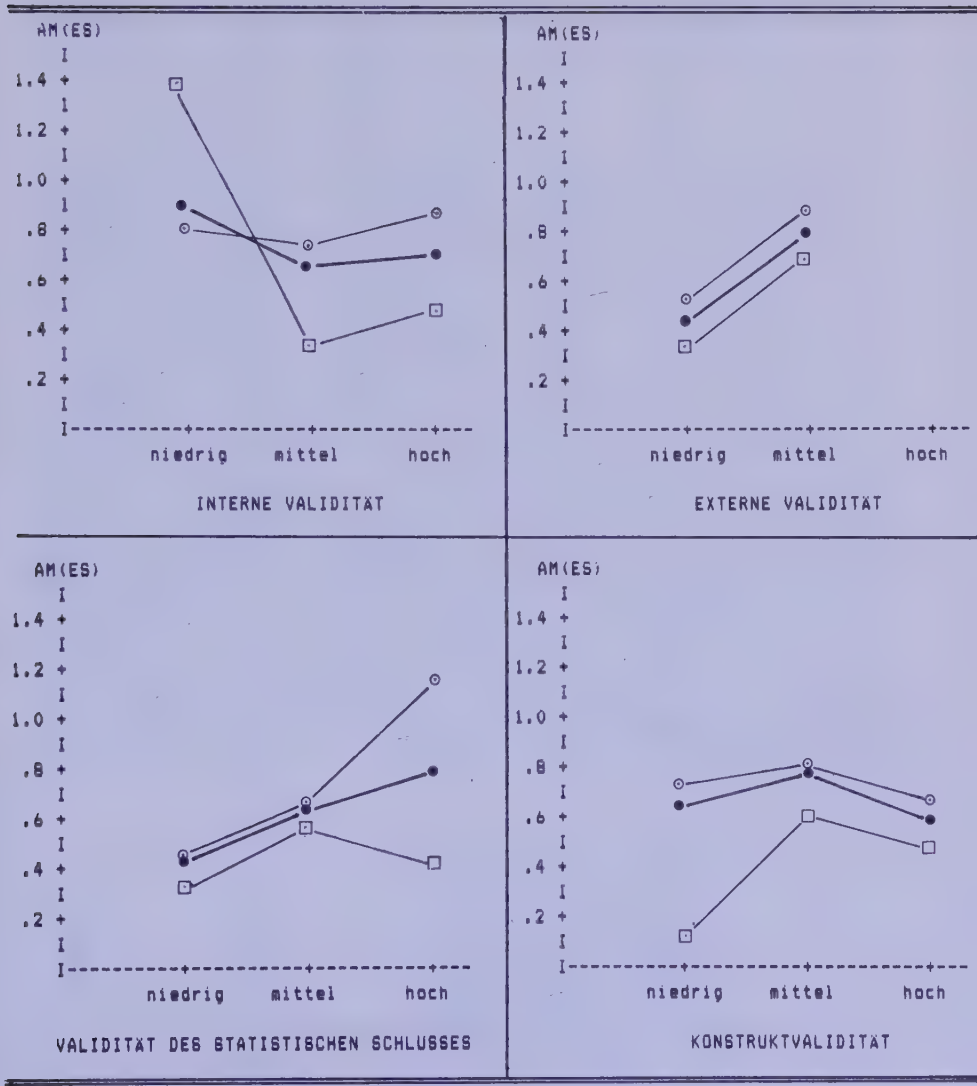


Die Abbildung zeigt, daß die Effekte für männliche Erwachsene höher liegen als für männliche Kinder. Umgekehrt zeigen weibliche Kinder einen Überalles Effekt, der über dem der weiblichen Erwachsenen liegt.

VALIDITÄT

Zwischen Studienvaliditäten und Effektgrößen besteht in Abhängigkeit des Validitätstyps ein Zusammenhang. Abbildung 33 zeigt besonders den Einfluß der Validität des statistischen Schlusses und den der externen Validität auf. Während für Zeitschriften-Beiträge ein linearer Zusammenhang zwischen Höhe der Validität des statistischen Schlusses und der Effektgröße besteht, ist für Disserationen der Effekt dann besonders niedrig, wenn die Validität des statistischen Schlusses hoch ist. Die Konstruktvalidität zeigt einen leichten und negativen Zusammenhang zum Effekt an. Besonders auffallend ist der Zusammehang zwischen niedriger Konstruktvalidität und niedrigem Effekt bei Dissertationen. Die interne Validität beeinflußt die Effekthöhe nicht. Für Zeitschriften und Dissertationen besteht ein Zusammenhang zwischen Effekt und externer Validität, wobei für keine Studie eine hohe externe Gültigkeit kodiert werden konnte.

Abb. 33: Effektgrößen in Abhängigkeit verschiedener Validitätsaspekte



5.2. Heuristische Varianzerklärungen

Der vorhergehende Abschnitt beschrieb die gemittelten Effektgrößen in Abhängigkeit verschiedener Merkmale. Im nachfolgenden Teil soll versucht werden, die Varianz der Effekte aufzuklären. Aufgrund der Datenqualität können diese Aussagen jedoch nur als heuristische Hinweise interpretiert werden.

Als varianzerzeugende Variablen werden solche Einflußfaktoren unterstellt, die sich aus den unterlegten Theorien bzw. Hypothesen wie aus den Hintergrundtheorien ableiten lassen. Abbildung 34 zeigt das unterstellte Modell. Jeder Effektgröße liegt eine Hypothese zugrunde, die die Referenz zu einer der Theorien zur Erklärung von Kooperation aufweist. Für jeden Effekt wird ferner unterstellt, daß die ihn erklärende Theorie 'wahr' sei und daß die zu untersuchende Hypothese korrekt aus dieser Theorie deduziert wurde. Sofern ceteris paribus-Klauseln nicht explizit genannt wurden, sie jedoch aus dem Kontext der Darstellung entnommen werden konnten, wurden sie berücksichtigt. Die Hypothesen beinhalten Aussagen über die Pradiktorvariable(n) und über die Kriteriumsvariable. Solche Aussagen können ggf. weitere theoretische Annahmen beinhalten (so wurde in frühen Untersuchungen die Kooperation als Ganzheit betrachtet, während im Rahmen der funktionalen Analyse sie in Einzelreaktionen zerlegbar ist). Die Anlage (Design) der Untersuchung, auf die sich der Effekt bezieht, stellt eine weitere mögliche Einflußquelle dar. Die Methode zur Gewinnung der Stichprobe können das Ergebnis ebenso beeinflussen wie die Merkmale der Probanden. Das methodische Vorgehen bei der Applikation des Treatments wirkt mutmaßlich auf das Kriterium. Neben Experimentator-Einflüssen, situativen Einflüssen und der 'subjektiven Echtheit' des Treatments (vgl. AUSUBEL 1983) wird hier insbesondere die individuelle Ausgestaltung des Meßinstruments als mögliche Einflußquelle unterstellt. Einen weiteren möglichen Einfluß kann das Auswertungsverfahren ausüben. Nicht nur die Validität des statistischen Schlusses, auch die Annahmen über die Datenqualität bilden potentielle Wirkbedingungen. Letztendlich erzeugt die Schätzung der Effektgrößen selbst mögliche Varianz.

Abb. 34: Deduktion verschiedener Einflüsse auf Effektgrößen

A) Suche nach einem Vorhersagemodell

Die vorangehende Analyse machte schon deutlich, daß die Effektgrößen sich sehr stark unterscheiden. Um hier nicht Extremwerte (sog. Ausreißer) dominieren zu lassen, wurden die Effektgrößen sowohl logarithmisch transformiert (vgl. WINER 1962, S. 397ff) als auch windsorisiert (vgl. DIXON & MASSEY, 1969). Selbst mit verschiedensten Gewichten versehene Regressoren ergaben keine bedeutsam höhere Varianzaufklärung, d.h., höhere multiple Korrelationen, als die Rohwerte. Allein diese Beobachtung führt zu dem Entschluß, die Effektgrößen so zu verwenden, wie sie vorlagen.

B) Selektion wesentlicher Variablen

Im nächsten Schritt galt es, aus der Vielfalt der Variablen jene auszuwählen, deren Einfluß auf die Effektgröße angezeigt werden konnte. Grundlage hierfür bildete eine Interkorrelationsmatrix. Die wesentlichen Prädiktoren wurden ausgewählt und schrittweise in die Regression eingegeben (vgl. KERLINGER & PEDHAZUR 1973). Tabelle 59 zeigt die varianzerklärenden Komponenten sowohl für alle Effekte wie in Abhängigkeit des Typs des Meßinstrumentes und in Abhängigkeit des Alters. Diese Variable wurde in 'Erwachsene' versus 'Kinder' geblockt, weil in der Literatur verschiedentlich darauf hingewiesen wird, daß die niedrigen Kooperationsraten Erwachsener auf die kulturspezifische Erfahrung zur Kompetition zurückgehe. Die Tabelle bildet nur solche Daten ab, deren Aufklärungsbeitrag groß oder signifikant ist. Die Übersicht zeigt, daß bis auf die Ausnahme 'alle Kinder' treatmentbezogene Komponenten zum größten Teil der aufgeklärten Varianz beitragen. Sie erklären durchschnittlich 44% der aufgeklärten Varianz. Für Experimente, die als Meßinstrument Nicht-Nullsummenspiele verwenden, liegt der Aufklärungsbeitrag bei 63%; bei Experimenten, die Handlungsverfahren wählten, hingegen bei 28%.

Tab. 58: Der Einfluß untersuchungsbedingter Komponenten auf die Effektgrößen

	Step 1	2	3	4	5	6	7	8	9	10	11	12	max. SUM
A VAR	UV*	Pj	D	SES	Alt	PA	Ges	SM	Val				
R**2	12	16	5	2	3	3	2	1	1				35
W VAR	UV*	Pj	Gz	Dt	Kp	Kk	Sm	Kd	Kf				
R**2	22	11	5	3	6	2	1	3	1				62
H VAR	UV*	Kk	Kf	Ges	In	Ae	Rk	SES	Pa	Rt	D	Kp	
R**2	9	8	6	3	2	2	2	2	1	3	1	2	45
E VAR	D	UVV	NK	Pa	UVN	SES	Su	Rk	T				
R**2	19	9	7	2	5	3	8	3	1				61
E VAR	D	UVV	NK	UVN	Rk	T	Su	Pa					
W R**2	22	12	8	5	4	5	2	1					64
E VAR	D	UVN	Kp	Pj									
H R**2	10	5	21	26									74
K VAR	T	UVN	D	MV	Val	Rt	Rk	Su	UVV				
R**2	10	4	4	5	3	2	4	2	1				42
K VAR	UVV	UVN	T										
W R**2	39	3	2										45
K VAR	UVN	D	MV	SES	Rk	Su							
H R**2	14	4	10	4	4	2							53

Legende zu Tabelle 56

A = alle Effektgrößen
W = Effektgrößen für Wahlmeßverfahren
H = Effektgrößen fur Handlungsmeßverfahren
E = Erwachsene
K = Kinder
VAR = Variablen
R**2 = Multipler Korrelationskoeffizient * 100
UV = unabhängige Variablen (aufsummiert)
UVV = unabhängige variablen (aufsummiert) vor dem Meßvorgang
UVN = unabhängige Variablen (aufsummiert) während des Meßvorgangs
Pj = Jahr der Publikation
D = Dauer der Treatmentapplikation
SES = sozioökonomischer Status
Alt = Alter der Probanden
Pa = Publikationsart
Ges = Geschlecht
Sm = Stichprobenmerkmale
Val = Validität
Gz = Gruppenzugehörigkeit
Ds = Schätzverfahren fur die Effektgröße
Kp = Kategorie der Problemstellung
Ks = Kalkulationsgrundlage der Effektgröße
Kd = Übereinstimmung zwischen Theorie und Hypothese
Kf = Kategorie der Fragestellung
In = Instruktion
Ae = Analyseeinheit
Rk = Referenzkonzept
Rt = Referenztheorie
Nk = Größe der Kotrollgruppe
Su = Stichprobenumfang
T = Transparenz der Untersuchung
MV = Meßinstrument der abhängigen Variablen

Die Tabelle zeigt zunächst, daß die transformierten Variablen - hier als aufsummierte Effekte UV, UVV und UVN - bzw. die Anzahl der Spielzüge, resp. die Dauer des Treatments den größten Varianzanteil (an aufklärbarer Varianz) erklären. Für Experimente, die Kooperation mittels PDG messen, werden durch die kodierten Variablen insgesamt 62% an Varianz zwischen den Effektgrößen aufgeklärt (Multipler Regressionskoeffizient (MRK) = .79); für Experimente, die Kooperation mit Hilfe von Handlungsverfahren messen, liegt der Aufklärungsbeitrag etwas niedriger (MRK = .67). Diese Zahlen belegen, daß unterschiedliche Studien, die dasselbe Phänomen untersuchen, dann integrierbar sind, wenn diese Unterschiede auf denselben Dimensionen (der transformierten Variablen) abgebildet werden können. Die Meta-Analyse liefert damit eine Möglichkeit, Effektunterschiede nicht nur auf der Basis von Bedingungen außerhalb der relevanten Untersuchungsvariablen beschreiben zu müssen, sondern diese relevanten Unterschiede selbst in den Mittelpunkt varianzaufklärender Verfahren zu stellen.

Der Einfluß der kodierten unabhängigen Variablen

Für die Übersichtsregression wurden die unabhängigen transformierten Variablen aggregiert. Dies geschah, um dieses Variablen-Set nicht übergewichtig in die Regression einfließen zu lassen und damit die Freiheitsgrade nicht überproportional durch diese Variablen determinieren zu lassen. Tabelle 59 zeigt den Einfluß der kodierten Variablen auf die Effektgrößen.

Tabelle 59: Schrittweise Regression der transformierten Variablen

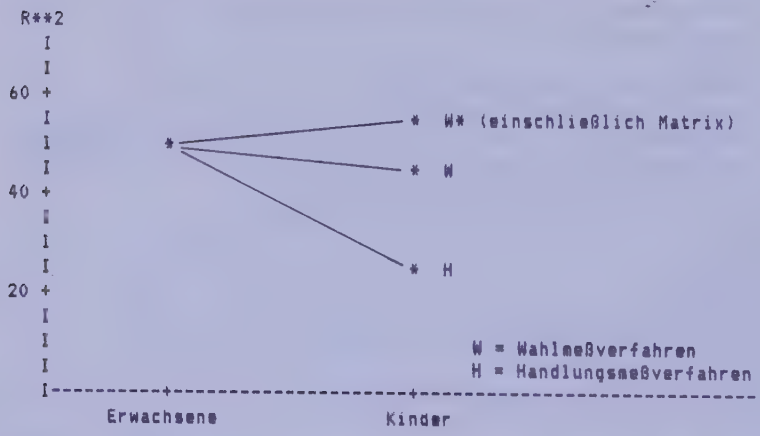
	Erklärende Variablen 1	2	3	4	5	SUM	N	AM(ES)	SD
A VAR	KBW	DHV	KBA	D					
R**2	14	4.5	2.1	1.		21.6	259	.65	.76
W VAR	DHV	D	DHW	REV					
R**2	21.8	20.3	5.	2.5		49.6	160	.61	.53
W VAR	DHV	D	DHW	REV	RE				
* R**2	21.8	20.3	5.	2.5	1.9	51.5	160	.61	.53
H VAR	KBW	DP	KBV						
R**2	19.7	6.5	3.1			29.3	99	.72	1.02
E VAR	D	DHV	DHW	KBW					
R**2	18.8	15.5	6.1	1.8		42.2	136	.64	.57
E VAR	D	DHV	DHW	REV					
W R**2	22	21.	5.3	1.8		50.1	136	.64	.54
E VAR	D	DHV	DHW	RE					
W* R**2	22	21.	5.3	3.1		51.4	136	.64	.54
E VAR	KBW	DP							
H R**2	19	31				50	20	.62	.76
K VAR	KBW	KBV	DHW						
R**2	17.6	2.5	4			24.1	103	.67	.98
K VAR	DHV	DHW							
W R**2	41.5	2.2				43.7	24	.43	.4
K VAR	DHV	S							
W* R**2	41.5	12.1				53.6	24	.43	.4
K VAR	KBW	KBV							
H R**2	22.5	4.1				26.6	79	.75	1.1

Legende zu Tabelle 59

A = alle Daten
E = Daten für Erwachsene
K = Effektgrößen für Kinder
W = Wahlverfahren
W* = Wahlverfahren unter Einbeziehung der Matrixwerte
H = Handlungsverfahren
VAR = Variable
R**2 = Determinationskoeffizienten * 100;
KBW = Kontingente Bekräftigung während des Meßvorgangs
KBA = Kontingente Bekräftigung alternativer Reaktionen vor dem Meßvorgang
KBV = Kontingente Bekräftigung kooperativer Reaktionen vor dem Meßvorgang
DHV = Diskriminative Hinweise vor dem MEßvorgang
DHW = Diskriminative Hinweise während des MEßvorgangs
D = Dauer
DP = Display
REV = Revenuen
RE = Reward (Zelle 1 in der PDG-MAtrix)
S = Sucker's payoff (Zelle 3 in der PDG-Matrix)

Abbildung 35 zeigt, daß die Varianzaufklärung für Erwachsene unabhängig des Meßverfahrens zu ca. 50% (MRK = .71) auf die kodierten Variablen zurückgeführt werden kann. Die Effektgrößen für solche Experimente, in denen Kinder als Probanden fungieren, werden in Abhängigkeit des Meßverfahrens unterschiedlich aufgeklärt.

Abb. 35: Varianzaufklärung in Abhängigkeit der Population und der Meßverfahren



In Experimenten der Handlungsmeßverfahren mit Erwachsenen übt die Treatmentdauer einen entscheidenden Einfluß aus. Ihr folgen diskriminative Hinweise vor und während der kooperativen Interaktion. Demgegenüber verhalten sich Kinder in Abhängigkeit der diskriminativen Hinweise vor der Interaktion und in Abhängigkeit des Verlustes bei nicht-erwiderter Kooperation (Matrixindex S). Für handlungsbezogene Meßverfahren liefern - unabhängig vom Alter - die Bekräftigungskontingenzen während der Interaktion einen wesentlichen Aufklärungsbeitrag und erst in zweiter Linie solche Treatmentbedingungen, die vor der Interaktion implementiert bzw. appliziert werden. Die Tabellen 60 - 63 weisen die interaktiven Beziehungen der signifikanten Prädiktoren aus.

Tab. 60: Prädiktoren für Wahlverfahren bei Experimenten mit Erwachsenen

Variable	(e)R**2	(c)R**2	sr	r	beta	
D (1)	22	22	.47	.47	.46	r(12) = .006
DHV (2)	21	43	.46	.46	.39	r(13) = .007
DHW (3)	5.3	48.3	.23	.40	.2	r(14) = -.004
REV (4)	1.8	50.1	-.13	-.00	-.16	r(23) = .34
						r(24) = .06
						r(34) = .39
RE (4')	3.1	51.4	-.18	-.34	-.19	r(14')= -.28
						r(24')= .01
						r(34')= -.23

(e)R**2 = quadrierte Semipartialkorrelation * 100
(c)R**2 = kummulierte Semipartialkorrelation * 100
sr = Semipartialkoeffizient
r = (Pearson-) Korrelationskoeffizient
beta = geschätztes Gewicht der Variablen

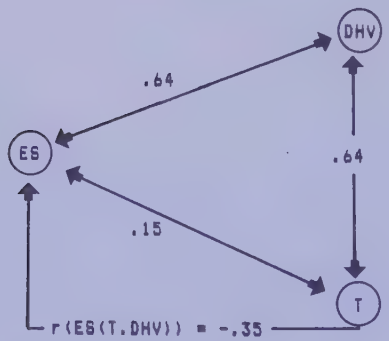
Tabelle 60 zeigt, daß die Effektgrößen für Experimente auf der Basis von Wahlmeßverfahren mit Erwachsenen wesentlich durch die Anzahl der Spielzüge und durch diskriminative Hinweise VOR dem Meßvorgang beeinflußt werden. Zwischen beiden Variablen besteht in der vorliegenden Untersuchung kein Zusammenhang. Bezieht man alternativ in die Regression statt der Revenuen den Matrixwert Re ein, kann ein Zusammenhang zwischen Re und Treatmentdauer als Anzahl der Spielzüge aufgezeigt werden: Je länger die Spielfolge, desto geringer fällt der Re-Wert aus. Dabei muß in Betracht gezogen werden, daß aufgrund der Beziehung T gt Re auch die Versuchungsvariable die Kooperationsbereitschaft beeinflußt (r(Re,T) = .78), wobei die Versuchungsvariable mit der Effektgröße zu r = -.34 korreliert. Damit wird dem Matrixindex für Kooperation immer ein adäquater Wert für Kompetition gegenübergestellt, um die Konfliktstruktur aufrechtzuerhalten. Bei instabiler Kooperation innerhalb der Episoden bedeutet dies ein hohes Verlustrisiko. Kontingenter Bekräftigungen während der Interaktion kommt nach dieser Analyse keine hohe Bedeutung zu. Wird die Zugfolge (Dauer) auspartialisiert, sinkt die Korrelation zwischen Effektgröße und KBW auf rp = .18. Demnach orientieren sich Probanden an der Kooperation ihrer Interakteure, die durch die Variable DHW kodiert wurde und weniger an den unmittelbaren 'absoluten Nutzen' in Form kontingenter Bekräftigung. Von Interesse ist der Zusammenhang zwischen der Form diskriminativer Hinweise während der Kooperation und der Revenuen. Experimente sind demnach so angelegt, daß ein positiver Zusammenhang zwischen beiden Prädiktoren besteht, jedoch die Revenuen auf den Effekt drükken.

Tab. 61: Prädiktoren für Wahlexperimente mit Kindern

Variable	(e)R**2	(c)R**2	sr	r	beta	Inter-Korrelationen
DHV (1)	41.5	41.5		.64	.56	r(12) = .45
DHW (2)	2.2	43.7	.15	.43	-.17	
T (2')	12.1	53.6	-.35	.15	-.35	r(12')= .64

Diskriminative Hinweise vor dem Meßvorgang bilden für Kinder eine wesentliche Bedingung für die nachfolgende Kooperation. Abbildung 36 zeigt darüber hinaus eine negative Supression:

Abb. 36: Supressionseffekt zwischen DHV und T



Wird die DHV-Variable aus T herauspartialisiert, kann der negative Zusammenhang zwischen Temptation und Effektgröße deutlich werden. Dieses Ergebnis stimmt dann mit dem für Wahlexperimente mit Erwachsenen überein.

Tab. 62: Prädiktoren für Handlungsexperimente mit Erwachsenen

Variable	(e)R**2	(c)R**2	sr	r	beta	
KBW (1)	19.1	19.1	.44	.44	.96	r(12) = .69
DP (2)	31.0	50.1	-.56	-.10	-.76	

Tab. 63: Prädiktoren für Handlungsexperimente mit Kindern

Variable		(e)R**2	(c)R**2	sr	r	beta	
KBW	(1)	22.5	22.5		.48	.41	r(12) = .31
KBV	(2)	4.1	26.6	.20	.34	.21	

Altersunabhängig liefern die Bekräftigungskontingenzen während des Meßvorgangs den größten Prädiktionsbeitrag. Bei jungen Probanden muß dabei die Bekräftigungskontingenz vor dem Meßverfahren berücksichtigt werden.

Die Wirkung der kodierten Variablen im Einzelnen

Auf der Basis der integrierten Studien zeigt die Ausformung kooperativer Verhaltensweisen durch Modellernen nur einen geringen Effekt. Dennoch liefert diese Variable in Handlungsexperimenten einen als signifikant ausgewiesenen Aufklärungsbeitrag. Für Wahlverfahren kann ein Einfluß jedoch nur aufgezeigt werden, wenn diskriminative Hinweise zur Situation oder Person herauspartialisiert werden. Diskriminative Hinweise während des Meßvorgangs führen zu Suppressionseffekten. Wird DHW herauspartialisiert, steigt der Einfluß der Modellierung von r = .0 auf sr(ES(MOD.DHW)) = -.53. Es wirken demnach solche Modellierungen effektiver, die nicht eindeutig die kontingente Interaktion darstellen. Bei Handlungsverfahren wird das Gegenteil angezeigt. Die Modelldarstellung der Interaktionskontingenzen von Episoden führen zu höheren Effekten.

Tab. 64: Modellierung als Basis-Treatment

		(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
Handlungsverfahren AM(ES) = .33 N = 24							
MOD	(1)	15.6	15.6		.4	.44	r(12) = .19
DHW	(2)	6.7	22.3	-.26	-.18	-.26	- - - - - - - -
							r(ES,D)= -.21
							r(1D) = -.19
							r(2D) = .76
Wahlverfahren AM(ES) = .29 N = 18							
DHW	(1)	21.8	21.8	-	.47	1.16	r(12) = .79
MOD	(2)	28.7	50.5	-.53	.04	-.87	

Kontingente Bekräftigung vor dem Meßvorgang führt zu hohen Effekten. Probanden, die an Handlungsexperimenten teilnahmen, zeigen auf der Basis dieser Variablen einen doppelt so hohen Effekt wie Probanden, die an Wahlexperimenten teilnehmen.

Tab. 65: Kontingente Bekräftigung vor der Interaktion als Basistreatment

Variable	(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
	Handlungsverfahren AM(ES) = 1.64 N = 12					
D	12	12	-	-.35	-.65	
REV	16.8	28.8	-.41	-.35	-.50	
	Wahlverfahren AM(ES) = .84 N = 22					
REV (1)	22.2	22.2	-	-.47	-.54	r(12) = .2
D (2)	10.6	32.8	.33	.23	.33	

Beide Datenmengen werden nicht durch die Variation der Variablen KBV selbst, sondern durch die Treatmentdauer und die Revenuen aufgeklärt. Dies ist insofern verständlich, als KBV kaum variiert. Revenuen sind bei beiden Meßverfahren für diese Teildaten von Bedeutung. Je ernster die Situation hinsichtlich möglicher Kosten ist, desto geringer ist die Effektgröße. Unterschiedlich beeinflußt die Dauer des Experiments die Kooperation. Während sie in Handlungsverfahren negativ korreliert, bestätigt sich in Wahlverfahren ihr unterstützender Einfluß.

Tab. 66: Kontingente Bekräftigung alternativer Verhaltensweisen als Basis-Treatment

Variable	(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
	Handlungsverfahren AM(ES) = .26 N = 6					
KBA (1)	64	64	-	.8	.8	
oder						r(11') = - 1.0
D (1')	64	64	-	-.8	-.8	
	Wahlverfahren AM(ES) = .99 N = 13					
DHV (1)	73.4	73.4	-	.86	.61	
KBW (2)	9.	82.4	.3	.77	.38	r(12) = .63

Werden während des Meßvorgangs alternative Reaktionen bekraftigt, liefern in Experimenten mit Wahlverfahren die diskriminativen Stimuli während des Meßvorgangs einen entscheidenden Einfluß sowie die während des Meßvorgangs vorherrschenden Kontingenzen. Die Erfahrung, daß auch mögliche alternative Reaktionen belohnbar sind, drücken im Rahmen der Untersuchungen nicht auf den Effekt. In Handlungsverfahren, denen als Treatment die alternative Belohnungsstruktur vorgelagert ist, wirkt diese Variable entscheidend auf das anschließende Verhalten: je intensiver alternative Reaktionen belohnt werden, desto niedriger fällt der Effekt aus. Für solche Experimente scheint aber auch - im Gegensatz zur über-alles-Betrachtung der Handlungsexperimente - die Dauer des Treatments und des Meßverfahrens von Bedeutung zu sein. Die kleine Anzahl verfügbarer Effektgrößen erlaubt an dieser Stelle jedoch keine weitere Interpretation.

Diskriminative Hinweise vor dem Meßvorgang wirken in Abhangigkeit des Meßverfahrens in Zusammenhang mit weiteren Variablen unterschiedlich. Die verschiedenen Niveaus der Variablen DHV selbst liefern keinen Aufklarungsbeitrag. Probanden, die an Experimenten teilnehmen, die in der Ankundigung zur Situation eine hinreichende Bedingung zur Ausformung kooperativen Verhaltens vermuten, stehen in ihrer Wirkung hauptsachlich unter dem Einfluß der Bekräftigungkontingenzen während der Interaktion.

Tab. 67: Diskriminative Hinweise vor dem Meßvorgang als Basistreatment

Variable		(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
		Handlungsmessung		AM(ES) = .28		N = 17	
KBW	(1)	60.4	60.4	-	.77	.63	r(12) = .39
DHV	(2)	11.0	71.4	.33	.61	.36	
		Wahlmessung		AM(ES) = .91		N = 27	
KBW	(1)	53.6	53.6	-	-.73	-.57	r(12) = -.41
D	(2)	12	65.6	.35	.62	.37	
T	(1')	57.4	57.4	-	-.76	-.64	r(1'2')=-.29
DHW	(2')	14.4	71.8	.38	.58	.40	

Während für Handlungsexperimente dieser Zusammenhang positiv angezeigt wird, liefern Wahlexperimente ein umgekehrtes Bild: Bei Experimenten, die als Basistreatment die Kooperativität der Situation angeben, sinkt die Effektgröße mit zunehmender Bekräftigung kooperativer Verhaltensweisen (vgl. die Diskussion dieser Anomalie). Eine mögliche Erklärung liefert die alternative Regression. Zwischen der Höhe des Versuchungswertes und dem Effekt besteht auch hier die mehrfach gefundene negative Beziehung. Versuchung und DHV korrelieren zu r = -.39. Je höher der Versuchungswert, desto niedriger zeigen vor dem Meßvorgang Informationen die Kooperativität der Situation an. Verstärkungskontingenzen können dann evt. nicht mehr eindeutig in ihrer Kontingenz erfahren werden.

Kontingente Bekräftigung während des Meßvorgangs führt in Abhängigkeit des Meßverfahrens zu unterschiedliche Effekten.

Tab. 68: Kontingente Bekräftigung als Basistreatment

Variable		(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
		Handlungsmessung	AM(ES) = .92		N = 39		
REV	(1)	19.6	19.6	-	-.44	-.40	r(12) = -.15
KBW	(2)	10	29.6	.32	.38	.32	
		Wahlverfahren	AM(ES) = .49		N = 66		
KBW	(1)	20.5	20.5	-	.45	.38	r(12) = .21
D	(2)	13.2	33.7	.36	.45	.37	

Die Übersicht zeigt, daß für Handlungsexperimente der Einfluß der kontingenten Bekräftigung während des Meßvorgangs wenig zur Aufklärung des Effektes beiträgt und zugleich eine negative Beziehung zwischen dem Niveau der Revenuen und den Effektgrößen aufgezeigt werden kann. Selbst bei Wahlverfahren klärt KBV als Basistreatment nicht den vermuteten hohen Varianzanteil auf. Da nun die Zugfolge für diese Teildaten bei AM(D) = 44 liegt, kann bei hoher Zugfolge (über 80% des Mittelwertes) und mittlerem Verstärkungniveau (AM = 3) ein Effekt von ES(schätz) = .94 erwartet werden.

Regressionen in Abhängigkeit der Referenztheorie

Wenn Theorien untschiedliche Hypothesen und damit Operationen zur Erklärung bzw. Manipulation des Phänomens Kooperation generieren, so müßten in Abhängigkeit der Theorien unterschiedliche Variablen die den Theorien zuordbaren Effekte erklären. Weiter kann angenommen werden, daß solche Variablen die Haupterklärung liefern, die in ihrer a posteriori-Klassifikation diese Theorien bzw. die aus ihnen deduzierten Hypothesen resp. Operationen am ehesten repräsentieren.

Für Untersuchungen, die sich auf das Bekräftigungsparadigma berufen, zeigt Tabelle 69 deutlich den wesentlichen Einfluß theoriebezogener Variablen.

Tab. 69: Regression der Effektgroßen in Referenz des Bekraftigungs-paradigmas

Variable		(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
		Handlungsmessung				N = 43	
KBV	(1)	31.3	31.3	-	.56	.53	r(12) = .49
KBA	(2)	10.3	41.6	-.39	-.38	-.32	
						.78 = a	
		Wahlmessung				N = 30	
PRG/S	(1)	21.1	21.1	-	.46	.54	r(12) =-.18
KBV	(2)	19.8	40.9	.5	.35	.45	
						.345 = a	

In diesen Experimenten wird ein Verhalten ausgeformt und seine Stabilität (bei unveränderten Kontingenzen) gemessen. Die Wirkung solcher Bekräftigungsstrukturen ist bei Wahlexperimenten vom Verhalten des Mitspielers abhängig. Spielt z.B. ein simultaner Ko-Akteur zu 80% kooperativ und erhielt der Proband eine Bekräftigungskontienz mittlerer Komplexität, so ist auf der Basis der vorliegenden Regression mit einem Effekt von ES(schätz) = 2.51 (= .78 + (.72 *(3)) - .73) zu rechnen.

Wie weiter vorne ausgewiesen, liefern Experimente auf der Grundlage des Paradigmas vom Modellernen keine eindeutigen Ergebnisse. Die nachfolgende Tabelle zeigt den Einfluß insbesonders diskriminativer Hinweise.

Tab. 70: Regressionen in Abhängikeit der Referenztheorie des Modellernens

Variable		(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
		Handlungsmessung		AM(ES) = .45		N = 27	
MOD	(1)	14.8	14.8	-	.39	.56	r(12) = .52
DHV	(2)	7.9	22.7	-.30	.04	- .33	
						.75 = a	
		Wahlmessung		AM(ES) = .35		N = 11	
DHV		36.7	36.7	-	.83		r(12) = -.38
MOD	(2)	12.5	49.2	-.45	-.56		

Die durch die Modellierung erzeugten Hinweise auf eine nicht-kooperative Situation führen zu höheren Effekten als die durch die Modellierung erzeugten Hinweise möglicher kooperativer Interaktionen. Ähnliche Zusammenhänge spiegeln auch die Ergebnisse innerhalb der Wahl-Experimente wider. Kooperationsanzeigende Modelle liefern eher Hinweise auf die (potentiell mögliche) Kompetitivität der Sitution (r(12) = -.38), wobei umgekehrt diskriminative Hinweise nicht unerheblich mit den Effektgrößen positiv korrelieren.

Austausch-theoretisch orientierte Untersuchungen zeigen ein recht homogenes Bild. Eindeutig zeigt die Kategorie der diskriminativen Stimuli ihren Einfluß auf die Effektgröße.

Tab. 71: Regression auf der Basis austauschtheoretischer Paradigmen
Wahlverfahren AM(ES) = .87 N = 23

Variable	(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
DHV (1)	57.1	57.1	-	.76	.69	r(12) = .2
KBW (2)	9.8	66.9	.48	.46	.32	
					-.19 = a	

Diese Beziehung ist mit austauschtheoretischen Annahmen konform: der Proband wird über seine Kosten-Nutzen-Möglichkeiten informiert und versucht, sich kosten-günstig zu verhalten. Kontingente Bekräftigung während des Meßvorgangs für das rationale, gebende Verhalten unterstützen das Aufkommen hoher Kooperations- bzw. Effektgrößen. Für diese Effekte gilt ein negativer Zusammenhang zwischen Effekt und Anzahl der Spielzüge (r = -.52): dieser Zusammenhang wird jedoch durch Auspartialisierung der Variable DHV eliminiert (r(ES(D.DHV)) = .04).: je kürzer die Anzahl der Spielzüge, desto weniger weisen in den Experimenten die Stimuli auf eine kooperative Situation hin (r(DHV.D) = -.71). Die Spielinstruktion korreliert für dieses Effekt-Sample ebenfalls mit der Effektgröße überdurchschnitlich hoch (r = -.72), d.h. die Spielinstruktionen enthält zunächst mit steigendem Effekt weniger Informationen über die rationale und kooperative Struktur des Spiels. Dies zeigt auch die hohe Korrelation zwischen DHV und Instruktion (r = -.84), die jedoch semipartiell auf sr(ES(INSTR.DHV)) = -.22 sinkt. Dies unterstellt, daß Experimente, die den Einfluß diskriminativer Hinweise als Treamtment variieren, den Einfluß über die Instruktion besonders deutlich kontrollieren wollen und primär mit dem Statement "mache den eigenen Gewinn" die Spiel-Regeln erläutern.

Experimente auf der Grundlage feldtheoretischer Paradigmen (sensu DEUTSCH) zeigen in Abhängigkeit des Meßinstrumentes unterschiedliche Einflußfaktoren.

Tab. 72: Regressionen auf der Basis feldtheoretischer Paradigmen

Variable	(e)R**2	(c)R**2	sr	r	beta	Interkorreltationen
	Handlungsverfahren		AM(ES) = .62		N = 24	
D* (1)	60	60	-	.77	.74	r(12) = -.66
DHV (2)	5.1	65.1	-.23	-.68	-.23	
					-1.82 = a	
	Wahlverfahren		AM(ES) = .59		N = 90	
KBW	32	32	-	.57	.46	r(12) = .29
DHV	12.3	44.3	.42	.50	.37	
					.06 = a	

Experimente im Rahmen der Handlungsmessung werden primär durch die Treatmentdauer in Minuten beeinflußt. Die diskriminativen Hinweise korrelieren mit der Effektgröße negativ, jedoch auch mit der Treatmentdauer. Lange Treatments liefern vor dem Meßvorgang kaum Hinweise zur Kooperativität der Situation. Zwar korreliert auch die Variable KBW mit dem Effekt (r = .75); sie sinkt jedoch bei Auspartialisierung der Variable D (r(ES(KBW.D) = .54). Theoretisch muß dieser Effekt umgekehrt erklärt werden: KBW liefert einen Erklärungsbeitrag von 56%, da die Variable DHV bei semipartieller Korrelation auf r(ES(DHV.KBW) = .12 fällt. D.h., kontingente Bekräftigung während des Meßvorgangs und diskriminative Hinweise vor dem Meßvorgang sind in den Experimenten hoch und negativ miteinander korreliert (r = -.81). Je eindeutiger Hinweise zur Kooperation auf interferente Fakrtoren hinweisen, desto höher unterstützt die Bekräftigungskontingenz während es Meßvorgangs Kooperation.

Für Wahlexperimente liefern dieselben Variablen den wesentlichen Erklärungsbeitrag. Nur besteht hier ein positver Zusammenhang zwischen KBW und DHV. Weitere Einflußfaktoren, wie z.B. die Instruktion (r(ES,INSTR.) = .49) oder die kontingente Bekräftigung alternativer Reaktionen vor dem Meßvorgang (r(ES,KBA)= -.55) sinken bei semipartieller Korrelation, so daß sie nicht weiter in der Regression behandelt wurden. Bezieht man tentativ die Bekräftigung alternativer Reaktionen ein (r(ES(KBA.KBW.DHV) = .30), so gilt: Für Wahlexperimente im Rahmen der feldtheoretischen Forschung wird die Effektgröße primär durch die Variablen KBW, DHV, KBA (zu 55%) erklärt. Für Effektgrößen im Rahmen dieses Forschungszweiges sind die Matrixindices von Bedeutung.

Tab. 73: Regression für feldtheoretische Paradigmen unter Berücksichtigung der Matrixindices

Variablen	(e)R**2	(c)R**2	sr	r	beta
Dauer	50.8	50.8	-	.71	.70
DHV	13.1	63.9	.36	.41	.24
P	3.5	67.4	-.18	-.18	-.20
DHW	1.8	69.2	.13	.40	.14
					.073 = a

Demnach übt die kontingente Bekräftigung während des Meßvorgangs keinen Einfluß aus (r(p)(ES.KBW.DHV.D.P) = .06). Wie im vorhergehenden Modell reduziert die Dauer bei Eintritt in die Regression durch ihre Interkorrelation mit der Variablen KBW (r = .47) deren Einfluß; auf nahe Null sinkt diese Wirkung jedoch erst bei Eintritt der Variablen P, obgleich KBW und P kaum miteinander korreliert sind (r = -.08). Ausschlaggebend dürfte die negative Korrelation zwischen DHV und P sein (r 0 -.4): je höher die P-Werte während des Spiels sind, desto weniger zeigen Stimuli vor dem Meßvorgang eine Kooperativtät der Situation an, so daß erst gar keine oder nur wenige kooperative Reaktionen ausgeführt werden, die zu einer Belohnung führen könnten.

Da die Feldtheorie mit umfangreichen Konstrukten operiert, wurden die Effekte nach (1) Instruktion zur Strategie, (2) Variation der Matrix-Indices und (3) Verhalten des Ko-Akteurs geblockt.

Tab. 74: Regression für die Variable Instruktion zur Strategie

Variable		(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
REV	(1)	29.1	29.1	-	.54	.81	r(12) = -.56
DHW	(2)	25.9	55	.6	.12	.60	r(13) = .17
DHV	(3)	12.9	67.9	.54	.47	.36	r(23) = -.06
						-.80 = a	

Entgegen der bisherigen Annahmen gilt für einen Teil der Daten (n = 18), daß Hinweise auf eine kooperative Situation und hohe Revenuen kooperationsfördernd wirken; dies setzt jedoch voraus, daß vor dem Meßvorgang auf die Interferenz der Situation verwiesen wurde.

Wird primär nach dem Einfluß von Matrixindices gefragt, zeigt die nachfolgende Tabelle, daß dieser Einfluß durch diskriminative Hinweise überdeckt ist.

Tab. 75: Regressionen für die Variable Variation der Matrixindices

Variable	(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
DHW (1)	54.2	54.2	-	.74	-.05	r(12) = .57
DHV (2)	8.4	62.6	.23	.66	.58	r(13) = .79
KBW (3)	9	71.6	.3	.65	.58	r(23) = .19
Intercept					.32 = a	

Obgleich die Variation der Matrixindices hier explizit untersucht wird, spielen sie für die Varianzaufklärung keine Rolle. Tabelle 76 zeigt die Korrelationen zwischen Matrixindices und Effektgrößen für dieses Sample.

Tab. 76: Interkorrelationen zwischen Matrixindices und Effektgröße

	RE	S	T	P	ES
RE	-	-.21	.83	-.36	-.50
S		-	-.38	.45	-.24
T			-	-.48	-.24
P				-	-.03

Der signifikante Einfluß des Reward-Index überrascht dabei. In die Regression der Tab. 75 wäre diese Variable an vierter Position in die Regression eingeflossen. Innerhalb dieses Datensamples und der sie repräsentierenden Untersuchungen gilt jedoch: dort, wo Zellenindies explizit untersucht werden und zusätzlich andere Variablen wirken (also kein Naturspiel), sinkt der Einfluß der Matrixindices ins statistisch Unbedeutende.

Die Frage nach der Wirkung über die Information zum Ko-Akteur weist darauf hin, daß diese Informaion erst durch die in vivo ermöglichte Verifikation von Bedeutung ist.

Tab. 77: Regression für Informationen zum Mitspieler

Variable	(e)R**2	(c)R**2	sr	r	beta	Interkorrelationen
DHW (1)	32	32	-	.57	.37	r(12) = .02
KBW (2)	23.3	55.3	.48	.49	.48	r(13) = .73
DHV (3)	3	58.3	.17	.54	.25	r(23) = .03
					-.088 = a	

Gegenüber den diskriminativen Hinweisen vor dem Meßvorgang wirken wesentlich die diskriminativen Hinweise während des Meßvorgangs und die erlebten Bekräftigungskontingenzen.

6. DISKUSSION DER ERGEBNISSE

Ziel der vorliegenden Meta-Analyse war es, eine Vielzahl von Untersuchungsergebnisse, die Kooperation auf unterschiedliche Art und Weise herzustellen beabsichtigten, in Abhängigkeit verschiedenster Merkmale zu vergleichen. Dieser Vergleich sollte trotz der breiten Variation der Treatments, Meßverfahren und Personenmerkmalen mit Hilfe transformierter Variablen eine Integration in dem Sinne erlauben, solche Bedingungen zu finden, die allgemein kooperationsfördernd wirken. Diesem Unterfangen lag die Vermutung zugrunde, daß ein und dasselbe Phänomen, nämlich kooperatives Verhalten, allgemein durch ein Netz von Variablen beeinflußt wird, unabhängig davon, welche theoretischen Terme diesen Variablen zugeordnet wurden. In die Meta-Analyse flossen sowohl Experimente ein, für die die jeweiligen Effektgrößen konventionell, d.h. im Sinne von standardisierten Mittelwertdifferenzen berechnet werden konnten, oder mit Hilfe eines Modells geschätzt werden mußten. Das hier speziell entwickelte Kodiersystem war anwendbar i.S. einer unterstellten Reliabilität und Validität, nachdem anhand von vier Untersuchungen (a) dieselbe Information kodiert wurde und (b) keine weitere Information, von der ein Einfluß auf das Ergebnis vermutet werden konnte, in den Studien gefunden wurde.

Entsprechend der entwickelten Tradition, in durchgeführten Meta-Analysen auch solche Variablen zu erheben, die keinen direkt ableitbaren theoretischen oder hintergrundtheoretischen Bezug aufweisen bzw. für die keine begründeten Hypothesen zur Verfügung stehen, wurden Publikationsmerkmale ebenfalls kodiert. Innerhalb des Erfassungszeitraumes zwischen 1956 und 1978, in dem die Untersuchungen publiziert wurden, steigt weder der mittlere Effekt noch sinkt die Varianz der Effektgrößen. Entgegen der Vermutung, daß im Verlauf eines Forschungsprozesses ein Phänomen zunehmend erklärt werden kann, zeigt Tabelle 30 zwar einen Anstieg der Effektgrößen zwischen 1961 und 1970; insgesamt aber kann anhand dieser Daten nicht davon gesprochen werden, daß im Forschungsverlauf konsistent die Evidenz von kooperationsfördernden Variablen systematisch entwickelt wurde; das indiziert sowohl der geringere Effekt in den nachfolgenden Jahren wie die Vergrößerung der Varianz. Obgleich ohne Ausnahme die Fragestellungen auf Möglichkeiten zur Reduzierung von 'self-defeating behavior' hinweisen, kann für Experimente auf der Basis von Wahlmeßverfahren vermutet werden, daß durch die konflikttheoretische Orientierung trotz positiver, d.h. Kooperation ausformender Bedingungen häufig Faktoren appliziert werden, die bei geringen Kooperationsraten zu Interferenzen führen (z.B. die Anzeige der 'unterschiedlichen' Auszahlungen). Tabelle 31 wirft ein Licht auf die Entwicklung verschiedener Treatmentvergaben. So nimmt zwar im Forschungsverlauf die Implementierung solcher Bekräftigungskontingenzen zu, von denen ein höher Effekt vermutet wird, auf der anderen Seite wird aber die Treatmentdauer zunehmend verringert - und dies in Experimenten, die komplexe Inter-Aktionen zu untersuchen beabsichtigen.

In Analogie zu den Ergebnissen von SMITH (1980) zeigen die vorliegenden Ergebnisse Verzerrungen in Abhängigkeit des Publikationstyps. Zeitschriften erzeugen höhere Effektgrößen (AM(ES) = .74; bei SMITH .64) als Dissertationen (AM(ES) = .43; bei SMITH .48). Tabelle 35 zeigt, daß in der vorliegenden Meta-Analyse Dissertationen sich an Referenztheorien orientieren, die zu niedrigeren Effekten führen. Im Rahmen des Modellierungsparadigmas versuchen Dissertationen Kooperation durch Reaktionsmodellierung eher auszuformen als durch Kontingenzmodellierung. Zudem werden in solchen Studien kürzere Treatments appliziert und jüngere Probanden untersucht.

Wesentliche Varianzquellen

Die allgemeine Vermutung, nach der die Unterschiede zwischen den Effektgrößen wesentlich auf die Unterschiede zwischen den Variablen, wie sie a posteriori kodiert wurden, erklärbar sind, kann in der Tendenz bestätigt werden. Die Zusammenfassung zeigt, daß diese Variablen in der Regel den größten Aufklärungsbeitrag liefern.

Tab. 78: Überblick varianzerklärender Variablen-Gruppen

Sub-Set	Variablengruppen					
Effektgrößen	TH-V	AV-V	VAL-V	UV	SONST	(c)R**2
			(R**2 * 100)			
alle Daten	-	-	.8	16.6	14.1	31.5
Wahlmessung	6.2	-	8.3	21.5	17	53
Handlungsmes.	7	-	2	12.3	21.8	43.1
Erwachsene	2.7	-	17.9	33.2	4.1	57.9
Erw./Wahl	4.1	-	14.5	39	.5	58.1
Erw./Handl	21	-	-	15	26	62
Kinder	6.2	4.8	14.4	9.3	-	34.7
Kind./Wahl	-	-	2.6	41.5	-	44.1
Kind./Handl.	3.7	9.1	2.2	18.4	4.9	38.3

TH-V = theoriegeleitete Variablen, AV-V = Bedingungen der abhängigen Variablen, VAL-V = Validitätsvariablen, UV = a posteriori kodierte Variablen, Erw. = Erwaschene, Kind. = Kinder, Wahl = Wahlmessung, Hand = Handlungsmessung

Für die aggregierten Daten unter der Kategorie 'Handlungsmessung', Erwachsene als Probanden bei der Handlungsmessung' und 'Kinder als Probanden liegt dieser Aufklärungsbeitrag jedoch sehr niedrig. Für Untersuchungen mit Kindern wird die Effektgröße durch die Validität der Studien stark beeinflußt, hier insbesondere durch das Item 'Transparenz der Fragestellung'. Die Tabelle zeigt und bestätigt vor diesem Hintergrund das GLASS'sche Argument, nach dem der Einfluß der Validität nicht a priori bestimmt werden.

sondern nur a posteriori erkannt werden kann. Die niedrige Aufklärung der Ergebnisse aus Experimenten zur Handlungsmessung kann u.U. auf Faktoren zurückgeführt werden, die nicht mitkodiert wurden. Dies deuten zwei Ergebnis-Tendenzen an:

(i) Kindern, denen unterstellt werden kann, auf freundliche Situationen sensibel zu reagieren, zeigen in Studien auf der Basis von Handlungsmessungen hohe Effektwerte, wie umgekehrt dort nur wenig Varianz auf die transformierten unabhängigen Variablen zurückführbar ist

(ii) auf der anderen Seite können solche Effekte besonders gut erklärt werden, die aus Experimenten unter dem Rational der Wahlmessung durchgeführt wurden und in denen die Treatmentapplikation hoch standardisiert ist.

Betrachtet man zusätzlich BERGMANN's Argument, nach dem sozialwissenschaftliche Theorien weder abgeschlossen noch vollständig sind, kann mit diesen Daten gleichsam ein weiters Argument von GLASS (1979) bestätigt werden:

"In none of the dozen or so research literatures that we have integrated in the past five years have we ever encountered a cross-validiated multiple correlation between study characteristics that was larger than approximately .60. That is, I haven't seen a body of literature in which we can account for much more than a t h i r d (kursiv im Original) of the variability in the r e s u l t s of s t u d i e s (kursiv im Original), ..." (S. 13)

Der 'subjektive' Faktor bei der Applikation von Treatments scheint in der Tat auch für Meta-Analysen von Bedeutung zu sein. So fordern z.B. AUSUBEL & THOL (1983) im Rahmen einer Diskussion um die durch Meta-Analysen evaluierte Wirksamkeit von Organisationshilfen beim Lernen (vgl. LUITEN et al. 1980; HORNKE 1983) die Ergebnisse durch Elemination "nicht-authentischer oder zumindest fragwürdiger Organisationshilfen..." (S. 65) zu homogenisieren. Diese Aussage ist insofern interessant, als das BARNES & CLAWSON (1975) zeigen, daß von 32 Studien zu 'advance organizer' 12 Untersuchungen die signifikante Überlegenheit dieser Lernstrukturhilfen anzeigen und 20 Untersuchungen keine signifikanten Unterschiede nachweisen können. Alle vier AUSUBEL-Untersuchungen sind jedoch signifikant. Interpretiert man AUSUBEL etwas frei, so heißt dies, Authentizität und damit Erkenntnis sind personengebunden. So verliert Forschung jegliche Rationalität im Sinne intersubjektiver Überprüfbarkeit. Auf der anderen Seite weist dieses Argument auf ein ernsthaftes Problem natürlich hin, daß für Datenintegrationen von Bedeutung ist. Die in den Publikationen kurz und prägnant beschriebe Vorgehensweise informiert nur unzureichend über Antezendenzen, die durch die Spezifikation der Treatmentvergabe zusätzlich wirken können. Damit sind nicht nur Experimentator-Probanden-Interaktionen gemeint, sondern auch die detaillierte Vergabe des Treatments in unstandardisierten Situationen. Wenn mit CRONBACH (1980) jede Treatmentvergabe als mögliche Realisation aus

einem Universum von Treatments bezeichnet werden kann, dann ist damit der Unterschied zwischen den Treatments festgeschrieben. Dies heißt jedoch, die Unterschiede genau zu erfassen, zumindest dann, wenn ihre Ergebnisse vom Durchschnitt der Ergbnisse abweichen. AUSUBEL wäre damit in die Pflicht genommen, seine Treatmentvergabe detailliert zu vermitteln.

Bei Differenzierung der transformierten unabhängigen Variablen kann dennoch die Varianz der Effektgrößen zu annähernd 50% aufgeklärt werden (vgl. Tabelle 59). Für Experimente im Rahmen der Wahlmessung liegt diese Aufklärung durchweg höher. Im Gegensatz zur Diskussion in der Literatur zeigen die Matrixindices einen nur unwesentlichen Beitrag auf. Auch dies indiziert die Brauchbarkeit der transformierten Variablen. Für Kinder scheinen demgegenüber die aus der Anlage der Spielregeln extrahierbaren - und in der Regel kontinuierlich abrufbaren - Informationen von handlungs- bzw. wahlentscheidendem Wert zu sein. Die Verlust-Variable S übt hier wesentlichen Einfluß aus: Kinder orientieren ihr aktuelles Verhalten nach dem eventuellen Verlust aufgrund einer 'Ausbeutungskontingenz', die dann eintritt, wenn die eigene kooperative Wahl nicht durch die ebenfalls kooperative Reaktion des Interaktionspartners erwidert wird. Kinder scheinen die Wahrscheinlichkeit einer Kooperation-Mogel-Kombination höher einzuschätzen als die einer Kooperation-Kooperation-Kombination. Je niedriger der mögliche Verlust ist, desto höher wird der Effekt.

Referenztheorien und Referenzkonstrukte

Die zentrale Vermutung, daß solche Experimente, die eine kooperative Interaktion indizieren oder deren Kontingenzen eine kooperative Inter-Aktion bekräftigen, höhere Effekte aufweisen, zeigt auch der Effektvergleich zwischen den Referenztheorien. Studien auf der Basis der operanten Konditionierung und austauschtheoretischer Annahmen zeigen gegenüber Experimenten auf der Grundlage der Feldtheorie sichtbar höhere Effekte. Diese Ergebnisse unterstützen die Annahme von McNEEL (1973), nach dem Bekräftigungskontingenzen suffiziente Bedingungen für kooperatives Verhalten darstellen. Er zeigt auf, daß 'own-gain-orientierte'(1) Probanden, von denen niedrige Kooperationsraten zu erwarten sind, sich dann kooperativ verhalten, wenn Kontingenzen Kooperation unterstützen. Die niedrigen Ergbnisse für Studien auf der Basis der Paradigmen des Modellernens überraschen. Eine

---

(1) In Analogie zu DEUTSCH (1949) wird in der Kooperationsliteratur zwischen verschiedenen Orientierungen differenziert. 'Own-gain'-orientierte Personen versuchen, unabhängig von der anderen Person ihr Ziel (bzw. ihren Gewinn) zu erzielen (individuelles Motiv); Demgegenüber beabsichtigen 'relative-gain'-orientierte Personen, durch kompetitives Verhalten (Motiv) den eigenen Gewinn auf Kosten des Interaktionspartners zu erhöhen. 'Joint-gain'-orientierte Personen sind bestrebt (motiviert), durch Kooperation einen gemeinsamen Gewinn zu erzielen.

weitergehene Analyse zeigt jedoch, daß diese Ergebnisse wichtige Hinweise für zukünftige Forschung liefern.

Innerhalb jeder Referenztheorie wurde zwischen verschiedenen Hypothesen und ihnen zuordbaren Konstrukten unterschieden. Die nachfolgende Übersicht zeigt die Effekte in Abhängigkeit dieser Konstrukte.

Tab. 79 : Effektgrößen in Abhängigkeit der Konstrukte verschiedener Theorien

Referenztheorie	Konstrukte	AM(ES)
Operante Konditionierung	Kontingente Bekräftigung kooperativen Verhaltens	.84
	Kontingente Bekräftigung alternativer Reaktionen	1.01
	Wahrnehmung der Bekräftigungskontingenz	1.25
	Modellierung	.37
Operante Konditionierung	Kontingente Bekräftigung kooperativen Verhaltens	.84
	Kontingente Bekräftigung alternativer Reaktionen	1.01
	Wahrnehmung der Bekräftigungskontingenz	1.25
	Modellierung	.37
Modellierung	der Kontingenz	.66
	kooperativer Reaktionen	.19
Austauschtheorie	Kontingente Bekräftigung kooperativen Verhaltens	.52
	Kontingente Bekräftigung alternativer Reaktionen	1.06
Feldtheorie	Kontingente Bekräftigung kooperativer Reaktionen	.69
	Vertrauen	.75
	Motivinduktion	.43

Die Tabelle zeigt Besonderheiten wie Widersprüche auf. So bezieht sich die Studie von MARTIN (1974) auf die operante Konditionierung und explizit auf

das Paradigma von SKINNER. Nur wird nicht die Bekräftigung kooperativer Verhaltensweisen in vivo, sondern durch ein Modell aufgezeigt. Die Effekte fallen noch niedriger aus als bei der Modellierung kooperativer Kontingenzen auf der Basis des Modellierungs-Paradigmas. Die Tabelle zeigt auch, das Untersuchungen im Rahmen der feldtheoretisch orientierten Kooperationsforschung nach DEUTSCH Kontingenzparadigmen angewandt werden (GUMPERT, DEUTSCH & EPSTEIN (1969), McCLINTOCK & McNEEL 1967). Sie zeigen niedrigere Effekte als operante Konditionierungen im Rahmen der Theorie SKINNERs. Die Vermutung liegt nahe, daß, wenn andere Hypthesen als die aus dem eigenen theoretischen Netzwerk, adaptiert werden, niedrigere Effekte eintreten. Dies kann an der weniger präzisen Umsetzung der entsprechenden Variablen liegen. Im Rahmen der Feldtheorie zeigen solche Untersuchungen, die Vertrauen zu induzieren beabsichtigen, innerhalb dieser Untermenge hohe Effekte. Die Feinanalyse weist für 56 Effektgrößen unter dem Konstrukt der Vertrauensinduktion eine hohe Präsenz von Bekräftigungskontingenzen und diskriminativen Hinweisen aus. Demnach besteht die methodische Umsetzung von Vertrauen darin, die Kooperativität der Interaktion bzw. des Interaktionspartners anzuzeigen und/oder solche Kontingenzen zu implementieren, die kooperatives Verhalten bekräftigen. Die diskriminative Information als Voraussetzung zur Kooperation diskutieren auch DEUTSCH (1958) und SOLOMON (1960). Demnach kann Vertrauen durch eine anfängliche Kommunikation über das Spielverhalten, also durch die Explikation von Strategien und deren Folgen, erzeugt werden. BIXENSTINE, POTASH, & WILSON (1963) versuchen, Vertrauen durch anfänglich hohe Kooperationsraten des simultanen Ko-Akteurs zu erzeugen. Die Ergebnisse weisen in dieser Form jedoch darauf hin, daß unkontingente hohe Kooperation Vertrauen darin fördert, den Ko-Akteur ausbeuten zu können; für unkontingent hohe Kooperationsraten ist die rationalste Strategie die der Mogel-Reaktionen; hierdurch sind die höchsten Gewinne zu erzielen. Dieselben Autoren zeigen auch, daß ethische Aspekte, z.B. den Ko-Akteur nicht ausbeuten zu wollen, eine untergeordnete Rolle spielen. Wahlverhalten scheint primär von austauschtheoretischen Aspekten gesteuert zu sein. LINDSKOLD & COLLINS (1978) untersuchen daher kontingente Koopertionsstrategien unter dem Vertrauensaspekt. Nach ihrer Ansicht beinhalten sowohl die GRIT-Strategie (1) wie die TFT-Strategie vertrauensevozierende Elemente. Demnach umfaßt daß Vertrauenskonzept nicht nur die antizipierte hohe Wahrscheinlichkeit, nicht ausgebeutet zu werden, sondern auch die antizipiert hohe Wahrscheinlichkeit, daß die eigene Ausbeutung bestraft

---

(1) GRIT (Graduated and Reciprocated Initiatives in Tension Reduction) stellt ein von OSGOOD (1962) entwickeltes Programm dar, das sowohl Vergeltung bei Ausbeutung wie unmittelbare Versöhnung beinhaltet. In der vorliegenden Untersuchung besteht die GRIT-Strategie aus folgenden Elementen:

(a) Ankündigung der Bereitschaft zur Konzilianz
(b) Exemplarische Durchführung konzilianter Reaktionen, d.h. unkontingenter kooperativer Reaktionen, mit dem Ziel, Vertrauen aufzubauen
(c) vergeltende Reaktionen auf Ausbeutungsverhalten nach der Einführungsphase.

wird. Vertrauen heißt dann die subjektiv kalkulierte Wahrscheinlichkeit einer Einhaltung der Reziprozität. Kontingente Spielstrategien stellen solche Mechanismen differentieller Bekraftigung dar. Kooperation des Akteurs wird mit Kooperation des Ko-Akteurs belohnt. Die Belohnung umfaßt dabei den hohen Punktwert für den Interaktionspartner und damit ggf. die negative Bekräftigung durch nicht-angewandte resp. erlebte Mogel-Reaktionen, verbunden mit dem diskriminativen Hinweis, unter bestimmten Bedingungen -nämlichen denen der Reziprozität - auch zukünftig kooperieren zu wollen. Zudem folgt die Strafe bei kontingenten Strategien auf dem Fuß. Fur kompetitives Verhalten wird der Partizipand durch den Ko-Akteur unmittelbar durch die kontingente Mogel-Reaktion bestraft.

Motivinduktion als kooperationsunterstützende Bedingung führt zu niedrigeren Effekten. DEUTSCH diskutiert drei grundlegende Motive, aus denen heraus Individuen eine Interaktion gestalten können: dem individuellen Motiv stellt er das kompetitive und das kooperative gegenüber. Die Aufgabe von Kooperationsexperimenten ist es nun, durch Variation der Antezendezen ein kooperatives Motiv zu induzieren. Explizit theoretische Konzeptionen zur Motivänderung werden nur marginal behandelt. SHAW (1958) differenziert kooperatives Verhalten hinsichtlich möglicher Antezendenzen in den motivationalen Faktor (als dem individuellen Bedürfnis, unter gegebenen Bedingungen etwas zu leisten) und dem prozeduralen Faktor (als Effektivität des sozialen Arrangement). Seine Ergebnisse zeigen auf, daß Probanden, die sich in einer kooperativen Situation 'glauben', hohe Performanzraten zeigen. Damit wird jedoch kein Hinweis auf eventuelle Motive geliefert.

Zusammenfassend kann aufgezeigt werden, daß die unterschiedlichen Theorien in ihrem Kern bestätigt werden können . Sie sind in der Lage, Kooperation zu erklären. Die Effekte in Abhängigkeit der Referenztheorien weisen auf der Basis der funktionalen Analyse jedoch darauf hin, daß (a) die sparsamen Theorien der operanten Konditionierung und der austauschtheoretischen Konzepte im Mittel die höchste Effekte erzeugen und (b) holistische Theorien auf der operationalen Ebene nicht ohne Elemente der sparsamen Theorien auskommen. Exemplarisch wird dies von IRLE (1975) herausgestellt.

"Drohungen und Versprechen sind V o r h e r s a g e n (im Original kursiv): P1 (oder P2) kündigen eine S t r a f e oder B e l o h n u n g an, die sie ausführen wird, wenn P2 (oder P1) innerhalb der Interaktions-MAtrix eine bestimmte Handlungsalternative anwenden wird" (S. 434).

Neben den vielfältigen Möglichkeiten, die Kooperation unterstützen, fördern oder Voraussetzungen bilden, haben sich in den Regressionen durchgängig drei Variablen herauskristallisiert:

(1) DISKRIMINATIVE HINWEISE, die über die potentielle Kooperativität der Interaktion informieren. Sie erleichtern die anschließende Rezeption der Kontingenzbedingungen

(2) KONTINGENTE BEKRÄFTIGUNG kooperativen Verhaltens während der Interaktion

(3) DAUER der Interaktion, in diesem Fall besonders die Anzahl der Spielzüge im Nicht-Nullsummen-Spiel.

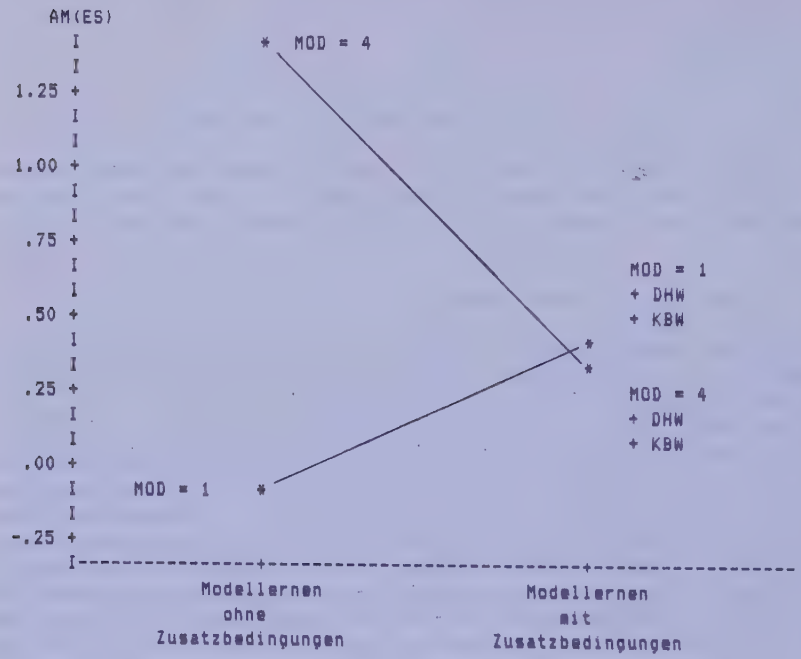
Interpretation der Forschungshypothesen

Während vorhergehend die Effektgrößen unter dem Aspekt der transformierten Variablen interpretiert wurden, werden in der Folge die Effekte auf der Basis der in der Literatur disktutierten Forschungshyothesen interpretiert.

(A) Das Paradigma des Modellernens

Verschiedene Studien (BRAVER & BARNETT, 1976, BRAVER & ROHRER, 1978, PAULSON, 1974) unterstellen, Modellernen sei eine hinreichende Bedingung zur Performanz kooperativen Verhaltens. Dabei wird angenommen, daß Probanden weniger die kooperative, sondern die 'günstigste' Kontingenz erwerben und im nachfolgenden Meßvorgang dieses Verhalten präsentieren. Wenn beispielweise ein Modell auf eine festgelegte, hohe Kooperationsrate des Ko-Akteurs mit Ausbeutung reagiert, so zeigt vermutlich auch der Beobachter bei gleichem Verhalten des Ko-Aktuers solche Ausbeutungsreaktionen. BRAVER & Van ROHRER (1978) vermuten die Überlegenheit der Verhaltensausformung durch Modellernen gegenüber unmittelbarer Kontingenzerfahrungen deshalb, weil letztere eine hohe Ich-Beteiligung involviert, die die Evaluation der Kontingenzstruktur erschwert. Tabelle 47 zeigt jedoch das Gegenteil an. Probanden, deren kooperatives Verhalten durch die modellierte Darstellung der Kontingenz ausgeformt werden sollte, zeigen einen geringeren Effekt als Probanden, die die Kontingenz unmittelbar erlebten. Abbildung 16 weist desweiteren die geringe Wirkung unmittelbar nachfolgender Bekräftigungskontingenzen und (hierdurch) vermittelter diskriminativer Hinweise zur Kooperativität aus, wenn die Probanden zuvor modellierte Kooperation sahen. Abbildung 37 stellt die Wechselwirkung verschiedener Variablen heraus:

Abb. 37: Modellwirkung in Abhängigkeit verschiedener Modellinhalte unter Berücksichtigung weiterer Antezendenzen



Zeigen Modelle einen niedrigen Anteil kooperativer Reaktionen (MOD = 1), so zeigen Probanden in der Performanz kein kooperatives Verhalten. Beobachten jedoch Probanden diese Modelle und sind sie selbst anschließend Kooperationskontingenzen ausgesetzt, die primär kooperatives Verhalten individuell bekräftigen, oder zumindest nicht-kooperatives Verhalten geringer bekräftigen als kooperatives Verhalten, zeigen diese Probanden Kooperation (AM(ES) = .47s). Wird hingegen eine individuelle Reaktionskontingenz belohnt und werden während des Meßvorgangs keine weiteren Variablen eingeführt, zeigen Probanden einen höheren Kooperationswert. Diese Rate sinkt unter der Bedingung, wenn solche Informationen angeboten werden, die das unkontingente Verhalten des Ko-Aktuers anzeigen. D.h., die niedrige über-alles-Wirkung der Modellvariablen wird durch Variablen während des Meßvorgangs hervorgerufen. Für Modellierungsexperimente im Rahmen der Wahlmessung gilt zusätzlich, daß mit zunehmender Ausführlichkeit der modellierten Kontingenzstrukturen wenig kooperatives Verhalten gezeigt wird, wenn zusätzlich der Proband den Ko-Aktuer als unkontingent und hoch kooperierend wahrnimmt. Die

Modelldarbietung informierte nicht nur über die Regeln der Kooperation, sondern auch über die Regeln der Ausbeutung.

Diese Ergebnisse stimmen mit der Untersuchung von BRAVER & BARNETT (1976) überein, die in ihren Experimenten widersprüchliche Ergebisse fanden. Probanden beobachten hier verschiedene kooperative bzw. kompetitive Interaktionen:

(1) beide Modellakteure kooperieren	KOOP(I)	KOOP(II)
(2) ein Akteur reagiert auf Kooperation des Ko-Aktuers mit Kompetition	KOMP(I)	KOOP(II)
(3) der Ko-Akteur reagiert mit Kompetition auf die Kooperation des Akteurs	KOOP(I)	KOMP(II)
(4) beide Modelle kompetitieren	KOMP(I)	KOMP(II)

Die Probanden spielten anschließend entweder gegen den Ko-Akteur oder gegen einen Fremden. Diese Ko-Akteure spielten grundsätzlich eine TFT-Strategie. Anzunehmen wäre, daß die Beobachter nicht nur die Kooperationsregel lernten, sondern auch die Ausbeutungsregel als sinnvollste Strategie bei unkontingenter und hoher Kooperation. Die Ergebnisse weisen dies jedoch nicht eindeutig nach.

Tab. 80: Ergebnisse aus der Untersuchung von BRAVER & BARNETT (1978)

Modellverhalten		Performanzbedingung	Prozent kooperativer Wahlen	
KOOP(I)	KOOP(II)	Modellpartner II	69	(1)
		Fremder	43	(2)
KOMP(I)	KOOP(II)	Modellpartner II	63	(3)
		Fremder	34	(4)
KOOP(I)	KOMP(II)	Modellpartner II	31	(5)
		Fremder	49	(6)
KOMP(I)	KOMP(II)	Modellpartner II	60	(7)
		Fremder	79	(8)

Demnach übernehmen Beobachter die Kooperationsregel, wenn sie unter denselben Bedingungen kooperieren können (1), und wenden diese Regel Fremden gegenüber verhaltener an (2). Überraschend ist das Ergebnis (3). Hier versucht der Beobachter nicht, den als ausbeutbar erlebten Modellpartner ebenfalls auszubeuten; vielmehr kooperiert er mit ihm. Von einem Fremden wird jedoch erwartet, dieser könnte einen selbst ausbeuten, so daß niedrige Reaktionsraten die Folge sind (4). Während die Ergebnisse (5) und (6) erwartungsgemäß ausfallen, weichen die Ergebnisse (7) und (8) von der Vermutung ab. Hier wurde nicht die kompetitive Strategie perzipiert, sondern die Unnützlichkeit dieser Strategie. Ähnlich argumentieren dann auch BRAVER &

BARNETT (1978). Sie gehen von der Annahme aus, daß die Kontingenzstruktur vom Beobachter gelernt wird und er zugleich annimmt, daß auch der ausgebeutete Ko-Akteur (2) sich im Wiederholungsfall nicht ausbeuten läßt. Dies machte zugleich die niedrigeren Kooperationsraten gegenüber dem fremden Ko-Akteur anstelle des kooperativen Ko-Aktuers für die Autoren deutlich. Von dem kompetitiven Modell-Ko-Akteur erwarten die Beobachter hingegen, daß er sich auch in Zukunft kompetitiv verhalten werde. In dieses Erklärungsmuster passen auch die letzten Ergebnisse. Die Beobachter perzipieren eine ir-rationale Strategie und bieten daraufhin eine kooperative an. Je weniger personale Identität zwischen Modell und aktuellem Ko-Aktuer besteht, desto kooperativer verhält sich der Beobachter: (8) gegenüber (7).

(B) Das Bekräftigungsparadigma

Zur Wirksamkeit operanter Konditionierung liegt ein ganzes Bündel an Hypothesen vor:

(1) Wenn kooperative Reaktionen kontingent bekräftigt werden, steigt die Wahrscheinlichkeit emittierter kooperativer Reaktionen

(2) Die Bestrafung und/oder die Antizipation von Strafkontingenzen für nicht-kooperative Reaktionen steigern die Auftrittswahrscheinlichkeiten kooperativer Reaktionen

(3) Kooperation ist dann höher auszuformen, wenn nicht nur kooperatives Verhalten belohnt, sondern zudem nicht-kooperatives Verhalten bestraft wird

(4) Hohes, unkontingentes Kooperationsverhalten durch den simultanen Ko-Akteur wirken - zumindest bei Wahl-Experimenten - als Belohnungskontingenz für kompetitives Verhalten. Hieraus leiten sowohl die Studien 1201, 1203 und 0204 die Hypothese ab, daß nicht-kontingente Kooperationsstrategien die Auftrittswahrscheinlichkeit kompetitiver Reaktionen erhöhen.

(5) Die niedrigen Kooperationswerte bei Wahl-Experimenten werden u.a. auf die niedrigen Revenuen , d.h. auf die ausschließliche Punktvergabe als 'pay-off' des Spiels zurückgeführt. Ein positiver Zusammenhang zwischen Revenuen und Kooperation wird vermutet (vgl. Studien 0201, 0701, 0803, 1201)

(6) Lehreraufmerksamkeit als Bekräftigungskontingenz ist eine hinreichende Bedingung zur Ausformung prosozialen - und darunter auch kooperativen Verhaltens (0202, 1901) .

(7) Nicht allein die applizierte Kontingenz, sondern auch das Wissen um die Kontingenz (awareness) erhöht die Kooperationsrate (2202, 2204).

Die Meta-Analyse zeigt auf, daß solche Treatments, die die kontingente Belohnung kooperativer Reaktionen vor dem Meßvorgang involvieren, zu einem mittleren Effekt von AM(ES) = 1.12s führen. Dieser Wert liegt nahezu doppelt so hoch wie für solche Untersuchungen, die als Treatmentbedingungen vor dem Meßvorgang keine Bekräftigungskontingenzen einsetzen. Die Ergebnisse unterstützen den von McNEEL (1973) bevorzugten sparsamen Ansatz zur Erklärung kooperativen Verhaltens, der auf intervenierende Variablen verzichtet. Für die Experimente, in denen die Variablen allein aus kontingenter Bekräftigung vor dem Meßvorgang besteht, d.h. die auf weitere Kontingenzen während des Meßvorgangs verzichten, liegt der Effekt bei AM(ES) = 1.1s. Er unterscheidet sich nur marginal von solchen Untersuchungen, in denen die Kontingenzen oder diskriminativen Hinweise während des Meßvorgangs aufrecht erhalten bleiben (AM(ES) = 1.2s). Gruppenkontingente Belohnung erzeugt gegenüber individueller Bekräftigung höhere Effekte:

$$(A)R(G) + (B)R(G) \longrightarrow 1\ (A,B)K(+)\ gt\ (A)R(G) + (B)R(G) \longrightarrow (A)K(+) + (B)K(+)$$

AM(ES) = 1.56s AM(ES) = 1.25s

Die Ergebnisse bestätigen auch die Wirkung kontingenter Bekräftigung während des Meßvorgangs. Für Experimente, die explizit oder implizit auf solche Kontingenzen während der kooperativen Interaktion verzichteten, beträgt der mittlere Effekt AM(ES) = .28s gegenüber solchen Experimenten, die Belohnungsstrukturen während des Meßvorgangs verankern (AM(ES) = .76s). Besteht die Möglichkeit, kooperatives Verhalten zu bestrafen, sinkt die Kooperationsrate unter die der Kontrollgruppe. Unerwartet fällt der Effekt für solche Experimente aus, in denen Kompetition höher belohnt werden kann als Kooperation.

Unkontingente, hohe Kooperation führt nicht unbedingt zu hohen Ausbeutungsreaktionen, wenn der Interaktion ein Treatment vorgelagert wurde. Die Ergebnisse von BRAVER & BARNETT (1978) indizierten dies bereits. Abbildung 27 zeigt dann auch eine leichte positive Korrelation zwischen unkontingenter Kooperationsrate und Effektgröße an. Diss bestätigen die Daten aus Tabelle 69. Für solche Experimente, die vor dem Meßvorgang ein Treatment einsetzten, daß Kooperation kontingent bekräftigt, korrelieren unkontingente Kooperation und Effektgröße zu r = .46.

Hohe Revenuen senken die Effektgröße. Dies bestätigen die Ergebnisse in den Tabellen 53, 60, 65, 68. Eine Ausnahme bilden solche Untersuchungen, in denen durch die Instruktion die Erwartung an eine kooperative Strategie

formuliert wurde (Tab. 74). Demnach wird erwartungskonformes Verhalten durch das Niveau der Revenuen unterstützt.

Soziale Verstärkung in Form kontingenter Lehreraufmerksamkeit führt zur Ausformung prosozialer Verhaltensweisen und hier auch zur Ausformung von Kooperation (AM(ES) = .89s). Die Effektgröße liegt unter dem Durchschnitt der durch kontingente Bekräftigung ausgeformten Kooperation. Dies mag darauf zurückgeführt werden, daß nicht SPEZIELL kooperative Interaktionen, sondern AUCH kooperative Interaktionen kontingent bekräftigt wurden.

(C) Das Diskriminationsparadigma

STAUB (1978, Bd. 1, S. 386f) verweist auf den Zusammenhang zwischen Stimuluskonfiguration und Kooperationsbereitschaft: je eindeutiger die Stimuli auf die Bereitschaft des Ko-Akteurs hinweisen, sich kooperativ verhalten zu wollen, desto höher verhält sich der Akteur kooperativ. Die Kooperativität der Interaktion kann auf zwei Ebenen angezeigt werden:

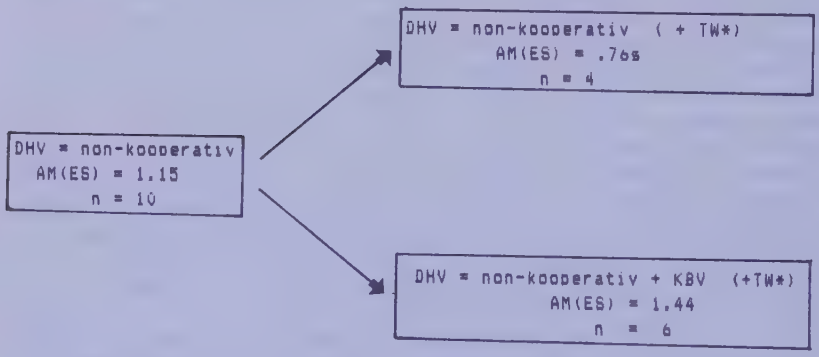
(a) durch Kommunikation (Instruktion) oder exemplarisches Verhalten vor dem Meßvorgang und/oder
(b) durch ein hohe Initialrate kooperativer Verhaltenssequenzen zu Beginn der Meßphase, d.h. der Interaktion.

Verschiedene Untersuchungen wenden diese Bedingungen an. So unterstellt die Untersuchung von McNEEL (1973) einen Zusammenhang zwischen der Instruktion und Kooperation. Den Gewinn-Aspekt betonende Instruktionen induzieren eine 'relativ-gain-Orientierung' und reduzieren die Kooperationsbereitschaft. Demgegenüber führen solche Informationen, die die Kooperationsbereitschaft des Ko-Akteurs anzeigen (GUMPERT, DEUTSCH, & EPSTEIN, 1969; LINDSKOLD, GAHAGAN, & TEDESCHI, 1970; JENSEN & MOORE, 1977) zu unterstellt hohen Kooperationswerten. OSKAMP & PERLMAN (1965) vermuten darüber hinaus, daß die Induktion einer 'Kooperationsnorm' gegenüber der Verankerung einer 'Kompetitionsnorm' die Kooperationsbereitschaft positiv beeinflußt. Solche Treatments sind häufig mit dem Konzept des Vertrauens von DEUTSCH verbunden. Danach verändert die Information, die eine Diskrimination kooperativer Strukturen ermöglicht, die subjektive Wahrscheinlichkeit für das Auftreten negativer Valenzen, d.h. solcher Reaktionen, die die eigene Kooperativität unterminieren. LINDSLEY (1966) sowie MARWELL, SCHMITT, & SHOTOLA (1971) nehmen daher an, daß Kommunikation per se risikolindernd wirkt und sich damit positiv auf die subjektive Wahrscheinlichkeit für negative Valenzen auswirkt. Nach SHAW (1958) reicht der 'Glaube', sich in einer kooperativen Situation zu befinden aus, um kooperativ zu reagieren. Umgekehrt sollen Probanden, die glauben, gegen einen kompetitiven Ko-Akteur zu spielen, einen niedrigen Kooperationswert aufzeigen. Die 'Glaubwürdigkeit' variiert die Studie von LINDSKOLD, GAHAGAN, & TEDESCHI (1970). Nach ihren Ergebnissen reicht der Glaube allein nicht aus, sondern Probanden müssen in der Lage sein können, durch ihre eigene Erfahrung die Situation bestätigen zu können, um höhere Kooperationsraten zu erzeugen. Die Anzeige des eigenen Verhaltens und die des Ko-Akteurs und der sich daraus ergebenden Auszahlungswerte (Display) gehören hierzu. McNEEL (1973) betont, daß die Anzeige der Differenz zwischen theoretisch möglichem und realem Punktwert zur Veränderung der Strategie hinreicht. Zuvor hatten McCLINTOCK & McNEEL (1966) sowie MINAS et al. (1960) angenommen, daß der Vergleich der Punktzahlen beider Ko-Akteure bei Differenzen die Kooperationsrate reduziere.

Der Einfluß diskriminativer Hinweise vor und während des Meßvorgangs weist kein homogenes Bild auf. Zum einen wirken die hierunter subsummierten Variablen häufig im Kontext weiterer Treatmentbedingungen, zum anderen erzeugen

sie auch dort widersprüchliche Ergebnisse, wo ihr Einfluß (qua Regression) 'dominant' ist. Tabelle 45 zeigt, daß Treatments, die die Information zur zukünftigen Interaktion einbeziehen, zu einem mittleren Effekt von AM(ES) = .84s führen. Abbildung 24 (Tabelle 50) weist demgegenüber aus, daß solche Bedingungen, die die Nicht-Kooperativität der Interaktion in Aussicht stellen, die höchsten Effektwerte liefern. Die nähere Datenbetrachtung weist 6 Effektgrößen aus, die neben der Variable DHV weitere Antezendenzen vor und während des Meßvorgangs realisieren:

Abb. 38: Die Wirkung diskriminativer Hinweise unter Berücksichtigung weiterer Bedingungen



* TW = Treatmentbedingungen während des Meßvorgangs

Ein Hinweis, der die Möglichkeit einer nicht-kooperativen Situation anzeigt, jedoch mit der kontingenten Bekräftigung kooperativen Verhaltens vor dem Meßvorgang verbunden ist, liefert dann hohe Effekte, wenn während des Meßvorgangs Tratmentbedingungen zur Unterstutzung der Kooperation verankert sind. Es ist zu vermuten, daß solche Hinweise als Ankündigung von Strafkontingenzen bei Nicht-Einhaltung der kooperativen Interaktion interpretiert werden. Das umgekehrt der Hinweis, die zukünftige Situation sei kooperativ, nicht ausreicht, um zu hohen Effekten zu führen, zeigt Abbildung 19. Liegt nur die Variable DHV vor und sind keine weiteren Bedingungen vor oder während des Meßvorgangs manipuliert, ergibt sich ein mittlerer Effekt von AM(ES) = .12s, während die Variable in Verbund weiterer Bedingungen zu einem Effekt von AM(ES) = .94s führt. Die Ergebnisse aus Tabelle 67 legen die Vermutung nahe, daß die diskriminativen Hinweise während der Interaktion evaluiert werden. Wenn auch JENSEN & MOORE (1977) zeigen konnen, daß bei Kindern eine kooperative Attribution ausreicht und zu höhren Effekten führt als eine kompetitive Attribution, bestätigen diese Daten doch eher

die Vermutungen von LINDSKOLD, GAHAGAN & TEDESCHI (1970), nach denen die Evaluation der Kontingenzen während der Interaktion verhaltensdeterminierend sind. WILSON (1971) variierte die Diskrimination der Situation, indem er in einem Vorspiel den simultanen Ko-Akteur entweder ganz kompetitiv oder im Verhältnis 3 : 1 kompetitiv spielen ließ. Im Verlauf der nachfolgenden 270 Züge, in denen dieser Ko-Akteur eine TFT-Strategie spielte, wählten die Probanden - zwar in Abhängigkeit von der Vorerfahrung, jdoch durchgängig - hoch kooperativ.

In der Tendenz stimmen diese Ergebnisse mit denen von McCLINTOCK & McNEEL (1966) überein. Probanden wurden hier vor einer Nicht-Nullsummen-Interaktion mit Simultanspielern konfrontiert, die sich entweder als 'gebend' (freundlich), als 'nehmend' (unfreundich) und neutral darstellten. Die Ergebnisse über 100 Züge zeigen zwar die Überlegenheit der freundlichen Bedingung gegenüber keiner Vorerfahrung. Jedoch zeigte auch die Gruppe der unfreundlichen Vorerfahrung eine Überlegenheit gegenüber keiner Vorerfahrung (F > UF > N). Für Probanden, die mit einem unfreundlichen Ko-Akteur konfrontiert waren, zeigen die Übergangswahrscheinlichkeiten (Stabilitätsmaß für die Aufrechterhaltung bestimmter Zugfolgen) dann eine hohe Stabilität, wenn die Inter-Akteure miteinander kooperierten. Dies läßt vermuten, daß solche Probanden, die zwar zuvor eine kompetitive Erfahrung machten, auf der anderen Seite jedoch die Möglichkeit zur Kooperation realisieren, die Strategie (erfolgreich) wechseln, wenn diese Strategie zu günstigen Kosten führt. Negative Verstärkung als der Entzug einer kompetitiven Strategie scheint hohe Kooperationsraten erzeugen zu können. Einen indirekten Hinweis liefern solche Beiträge, in denen über die Spiellänge diskutiert wird. In diesem Zusammenhang wird nicht selten die Vermutung geäußert, die Probanden hätten keine Zeit mehr, aus ihren negativen Erfahrungen zu lernen und durch Rückmeldung zu einer kooperativen Strategie zu wechseln.

In Untersuchungen, die durch Motiv-Induktion versuchen, Kooperation zu beeinflussen, wirkt die Variable DHV wesentlich (vgl. Tabelle 72). Motivinduktion wird sowohl bei SHAW (1958) wie auch bei GARNER & DEUTSCH (1974) als Vermittlung einer Handlungsorientierung durch Instruktion verstanden. Während SHAW unmittelbar durch die Begriffe 'Kooperation' und 'Kompetition' eine experimentelle Situation , in der sich die Probanden befinden, etikettieren, umschreiben GARNER & DEUTSCH die assoziierten Attribute und Verhaltensweisen der jeweiligen Situation bzw. Interaktion. Motivation wird hier also als durch Instruktion vermittelte Orientierung betrachtet (vgl. auch KELLEY & STAHELSKI, 1979). Die Ergebnisse von GARNER & DEUTSCH demonstrieren den interdependenten Charakter diskriminativer Hinweise und Bekräftigungskontingenzen. Probanden erhielten unterschiedliche Informationen, um eine kooperative gegenüber einer kompetitiven Orientierung zu induzieren. Anschließend wurden diese Probanden Mitspielern ausgesetzt, von denen sie annehmen mußten, diese hätten dieselbe Orientierung wie sie; in der Tat verfügten die Mitspieler jedoch uber eine entgegengesetzte Orientierung. Umgekehrt nahmen andere Probanden an, ihre Mitspieler verfügten über eine entgegengesetzte Orientierung; hier lag dann real dieselbe Orien-

tierung vor. Unter der Kontrollbedingung waren die Probanden nicht über die Orientierung ihrer Ko-Akteure informiert worden. Es interagierten also immer ein kooperativ-orientierter und ein kompetitiv-orientierter Spieler. Tabelle 81 gibt die Ergebnisse wieder.

Tab. 81: Prozentwerte kooperativer Wahlen im Experiment von GARNER & DEUTSCH (1974)

Experimental-Bedingung	Wahl-Blöcke	
	1 - 15	16 - 30
NICHT-ÜBEREINSTIMMENDE INFORMATION		
kooperative Probanden	61.3	50.0
Kompetitive Probanden	48.7	38.7
KEINE INFORMATION		
Kooperative Probanden	61.3	63.3
Kompetitive Probanden	48.0	48.0
ÜBEREINSTIMMENDE INFORMATION		
Kooperative Probanden	58.7	72.0
Kompetitive Probanden	43.3	57.3

Probanden, die die (falsche) Information erhielten, mit einem kooperativen Ko-Akteur zu interagieren, verhielten sich zu Beginn kooperativ und adjustierten im Verlauf der Zugentwicklung ihre Strategie auf den realen kompetitiven Ko-Akteur. Kooperativ-orientierte Pobanden, die über die Orientierung ihres Mitspielers nicht informiert worden waren, blieben in ihrer Kooperationsrate stabil. Kooperativ-orientierte Probanden, die valide über die Orientierung der Ko-Akteure informiert worden waren, erhöhten im Verlauf der Züge ihre Kooperation. Überraschend ist hier jedoch auch die Zunahme der kompetitiv orientierten Probanden, zumal 68% der kooperativ und 68% der kompetitiv orientierten Probanden das Spiel als kompetitiv charakterisierten. Kompetitiv orientierte Personen wurden als einflußreich bezüglich des Spielverlaufs eingeschätzt. Die Ergebnisse zeigen, daß selbst kompetitiv valide eingeschätzte Interaktionen durch Entwicklung einer Strategie kooperieren können, solange beide Interakteure in ihrer Verhaltenstruktur zwar interdependent, jedoch nicht dependent verbunden sind. Die insgesamt hohe Kooperationswerte in dieser Untersuchung sind durch die Matrixindices beeinflußt.

Die diskriminativen Hinweise während des Meßvorgangs üben einen starken Einfluß auf die Kooperationswerte aus. Je eindeutiger die Hinweise auf die Koopertivität der Interaktion hinweisen, desto größer fällt der gemittelte Effekt aus. Dennoch trägt diese Variable über alle Daten wenig zur Varianzaufklärung bei. Für Experimente im Rahmen der Wahlmessung korreliert sie

zwar sowohl für die Population der Kinder wie für die der Erwachsenen mit der Effektgröße zu r = .43 bzw. r = .40, ist jedoch zugleich mit der Variable 'Diskriminative Hinweise vor dem Meßvorgang' korreliert (r = .45 bzw. r = .34), so daß ihre semipartielle Korrelation auf sr = .15 rsp. sr = .23 sinkt. Unter anderen Selektionskriterien reduziert sich die Korrelation ebenfalls unter dem Einfluß weitere Variablen. Die Vermutung liegt daher nahe, daß diskriminative Hinweise während der Interaktion als solche von untergeordnetem Einfluß sind und erst in Zusammenwirken mit Informationen vor dem Meßvorgang (Selektion oder Hinlenken der Wahrnehmung, repräsentiert durch die Variable DHV) oder durch die Modellierung von kooperativen Interaktionen (vgl. Tabellen 64 und 67) von Bedeutung sind. Dies belegen exemplarisch die Ergebnisse aus der Tabelle 74 und 66. Die Effektgrößen, die auf den Einfluß der Instruktion zum Spiel zurückgehen, liefern nach Adjustierung um den Einfluß der Revenuen den wesentlichsten Varianzbeitrag (sr = .6). Modellierung und diskriminative Hinweise während des Meßvorgangs sind zu r = .79 korreliert; Hinweise während des Meßvorgangs klären jedoch zu einem größeren Teil die Abweichungen von der mittleren Effektgröße für dieses Daten-Sample auf (r = .47).

Kontinuierliche Hinweise zur Kooperativität bzw. Reziprozität der Interaktion liefert die Anzeige des Spielverhaltens bzw. -erfolges (Display). So lassen z.B. GUMPERT, DEUTSCH & EPSTEIN (1969) nach der Wiedergabe der Spielzüge und der entsprechenden Punkteverteilung die Akteure ihre Erwartung hinsichtlich des nächsten Zuges niederschreiben. Damit sind die Teilnehmer gezwungen, aus den diskriminativen Hinweisen, die sie pro Zug erhalten, ihr weiteres Verhalten abzuleiten. Bei einer - über alle Treatmentvariation in dieser Studie hinweg betrachteten - mittleren Kooperationsrate von 43% kognizieren die Probanden eine zumindest ambiguente Situation, die die zukünftige Reaktionswahrscheinlichkeit für kooperative Wahlen mindert. So ist es nicht verwunderlich, daß Experimente, in denen die Spielverläufe angezeigt werden, unter der Bedingung niedriger anfänglicher Kooperationsraten in der Tat auch zu niedrigen Effekten führen. Experimente, in denen versucht wurde, durch beginnend hohe Kooperationsquoten durch den simultanen Ko-Akteur Vertrauen zu induzieren (vgl u.a. BIXENSTINE, POTASH & WILSON, 1963), erlauben die Diskrimination, daß der Ko-Akteur die Möglichkeit zur Kooperation anbietet. WILSON (1971) zeigt, daß anschließend reaktionskontingentes Verhalten vor dem Hintergrund einer solchen Vor-Erfahrung zu hohen und stabilen Kooperationsraten führt. LINDSKOLD & COLLINS (1978) erweitern mit Hilfe der GRIT-Strategie das Ensamble möglicher diskriminativer Stimuli. Die GRIT-Strategie erlaubt einem Ko-Akteur, seinen Interaktionspartner über die seines Erachtens nützlichste Strategie zu informieren und bei Mißbrauch mit Vergeltung zu reagieren. GRIT-erfahrene Probanden perzipieren sehr schnell unkontingente und hohe Kooperationsraten, auf die sie mit Mogeln reagieren. Demgegenüber halten Probanden, die eine TFT-Strategie kennenlernten, ihre Kooperationsquoten nahezu aufrecht. Daraus kann tentativ geschlossen werden, daß Personen, die an einer kooperativen Interaktion interessiert sind, sowohl ihre Bereitschaft zur Kooperation (modellhaft) anzeigen müssen, wie sie weiter zu verdeutlichen haben, daß sowohl

eine kooperative Strategie effektiv ist und sie daher sich nicht ausbeuten lassen (in Analogie zur TFT-Strategie).

(D) Interferente Einflüsse

STAUB (1979) stellt nicht nur die 'eindeutige' Stimulus-Konfiguration als Voraussetzung für die Bereitschaft, sich kooperativ zu verhalten, heraus. Nach ihm reduzieren gleichsam solche Faktoren die Kooperation, die zu dieser 'Eindeutigkeit' interferent sind. Als interferente Bedingungen können aus der Sicht der Ergebnisse dieser Meta-Analyse solche Reaktionen herausgestellt werden, die das kooperative Verhalten der Probanden (durch Nicht-Kooperation wie durch externe Faktoren) bestrafen bzw. Mogelreaktionen belohnen. Die Betonung der Bereitschaft liegt darin, sich kontingent und nicht bedingungslos kooperativ zu verhalten. Einen weiteren, die Kooperation einschränkenden Faktor bildet die Lerngeschichte i.S. bereits etablierter Kontingenzen und Dikriminations- wie Generalisationsrepertoirs der Probanden. Die nicht in die Meta-Analyse aufgenommenen Untersuchungen von MADSEN (1967), KAGAN & MADSEN (1971) und SHAPIRO & MADSEN (1969) untersuchen den Einfluß unterschiedlicher Kulturen - hier insbesondere den Erziehungsstil - als mögliche Einflußquelle kooperativer Verhaltenswahrscheinlichkeiten. Dabei werden weitere Variablen, wie der sozioökonomische Status, auf dem Land oder in der Stadt aufgewachsen zu sein und das Geschlecht kontrolliert. Hierbei wird die Diskussion aufgegriffen, daß amerikanische Sozialisation zur Kompetition erzieht und daß Männer ökonomisch orientierter und also kompetitiver seien als amerikanische Frauen. Anzunehmen ist demnach, daß Frauen kooperativer wären als Männer und das Probanden, die auf dem Land groß geworden sind und aufgund der traditionellen Kooperation weniger kompetitiv orientiert wären, sich kooperativer verhielten als Männer bzw. als Probanden, die in einem städtischen Milieu groß wurden. Ob der sozio-ökonomische Status über die Leistungsortientierung mit der Kooperationsbereitschaft korreliert, ist eine weitere Frage. WHITE (1977) zeigt in seiner Meta-Analyse zwar einen Zusammenhang zwischen sozio-ökonomischem Status und Schulleistung; er weist gleichsam darauf hin, daß dieser Zusammenhang mit zunehmender Verweildauer in der Schule sinkt. Nach seiner Interpretation kann angenommen werden, daß die schulische Sozialisation erfahrungsegalisierend wirkt. Entgegen den Erwartungen auf der Grundlage spezifischer Untersuchungen zeigen männliche Erwachsene höhere Kooperationsraten als männliche Kinder und weibliche Probanden. Die höhere Kooperationsbereitschaft von Männern wird auch durch die Ergebnisse angezeigt, in denen zwar gleichgeschlechtliche Dyaden interagierten, die Daten jedoch geschlechtsunabhängig ausgewertet wurden. Die mittlere Effektgröße (AM(ES) = .42) liegt zwischen der für männliche Probanden (AM(ES) = .8) und ausschließlich weibliche Probanden (AM(ES) = .22).

(E) Besondere Variablen im Rahmen von Nicht-Nullsummenexperimenten

Innerhalb der Kooperationsforschung auf der Basis von Nicht-Nullsummenspielen werden drei Instrument-spezifische Merkmale diskutiert:

(i) Matrix- und Zellenindices
(ii) Spielstrategien der simultanen Ko-Akteure
(iii) Anzahl der Züge (Spiellänge)

Ein Spiel, bei dem nur die Matrixindizes und/oder die Anzahl der Züge kontrolliert bzw. variiert werden, wird als natürliches Spiel (natural play) bezeichnet. Hier lassen sich vermutete Einflüsse aufgrund der Spieldauer und der Zellenwerte genauer studieren.

(i) Matrix- und Zellenindices
Matrix- und Zellenindices kommt in Abhängigkeit davon, ob nur die Matrixwerte variiert werden (natürliches Spiel) oder ob weitere Treatmentbedingungen vorliegen, unterschiedliche Bedeutung zu. Tabelle 55 zeigt, daß im natürlichen Spiel der Bestrafungswert P einen hohen Zusammenhang zur Kooperation ausweist: je höher die mögliche Bestrafung für eine nicht-kooperative Interaktion ausfällt, desto kooperativer verhalten sich Probanden (r_s = -.55). Diese Tendenz bestätigt die Ergebnisse der Untersuchung von STEELE & TEDESCHI (1967). Sie finden für CC- und UL-Wahlen (1) eine Korrelation von -.31 bzw. -.35.
Damit kann der in der Literatur diskutierte Einfluß des P-Index bestätigt werden (vgl. McCLINTOCK & MESSICK, 1967, OSKAMP & PERLMAN, 1965). RAPOPORT & CHAMMAH (1966) zeigen dabei jedoch auf eine kurvi-lineare Beziehung hin: die Kooperationsraten steigen, wenn der P-Wert kleiner oder größer als -10 betrug, zunehmend. Werden neben den Matrixwerten weitere Variablen appliziert, verliert der P-Wert an Bedeutung: stattdessen beeinflussen der Re- und der T-Index das Verhalten mit. Je höher der Belohnungswert für Mogelreaktionen ist, desto niedriger fällt der Kooperationseffekt aus. Wenn Probanden - über alle Treatmentunterschiede hinweg - dahingehend orientiert werden sollen, sich kooperativ zu verhalten, dann wirkt jedoch nicht nur der Versuchungsindex T, sondern auch die Höhe des Verstärkungsindex Re kooperationsmindernd. Dies widerspricht der von RAPOPORT (1966) aufgezeigten Tendenz, daß kooperative Wahlen steigen, wenn Verstärkungsindex und Bestrafung für Kooperation steigen.

Um die Wechselwirkung der verschiedenen Matrixwerte kontrollieren zu können, wurden unterschiedliche Zellenindizes vorgeschlagen. Die Bereitschaft zur Kooperation ist demnach von der Differenz zwischen Belohnung für Kooperation (Re) und Versuchung, nicht zu kooperieren (T) abhängig. Diese Tendenz wird jedoch durch die Gefahr, der andere Mitspieler verfolge eine

---

(1) als CC-Wahlen werden bilaterale und als UL-Wahlen unilaterale Kooperationswahlen bezeichnet.

gegen den Akteur gerichtete Strategie konterkariert. Unter der Bedingung, kooperieren zu wollen, bildet daher die Differenz zwischen gelungener Kooperation (Re) und Verlust infolge mißlungener Kooperation (S) eine mutmaßlich ausschlaggebende Größe: ihre Differenz hemmt die Kooperation im Sinne eines wachsenden Risikos. Da nun für die Matrixindices die Beziehung T gt Re und Re gt S gilt, kann folgende Vermutung angenommen werden: Je größer der negative Re-T-Wert und je größer der Re-S-Wert, desto geringer ist die Kooperationsbereitschaft.

Ähnlich argumentiert RAPOPORT (1967), wenn er meint, daß die Bereitschaft, kooperativ zu wählen, durch die Differenz T - Re primar genemmt wird, während die Bereitschaft kompetitiv zu wählen durch die Differenz zwischen Re - P gestört wird. Tabelle 56 zeigt die Einflusse verschiedener, auf dieser Basis berechneter Zellenindices für alle Effektgrößen im Rahmen der Nicht-Nulsummenexperimente im vergleich zu den Daten aus Experimenten zum Naturspiel. Für das Naturspiel korrelieren die von RAPOPORT diskutierten Indices

$$r(1) = \frac{Re - P}{T - S} \quad \text{zu } r(s) = .48 \quad \text{und}$$

$$r(4) = \frac{T - Re}{T - P} \lg \quad \text{zu } r(s) = -.35$$

r(1) drückt die Gewinnmöglichkeit bei kooperativen Wahlen aus. Die Gewinne unter der Bedingung kooperativer Wahlen, wenn der Ko-Akteur kooperiert oder kompetitiert, werden der Gewinnmöglichkeit bei eigener Kompetition gegenübergestellt. Je höher dieser Index den Vorteil einer Kooperation ausdrückt, desto höher fällt die Kooperationsbereitschaft aus. r(4) drückt demgegenüber die Versuchung relativ zu ihrem Gelingen aus. Die Kooperationsbereitschaft wächst umgekehrt proportional mit diesem Quotienten. Für kompetitiv orientierte Probanden ist der P-Wert daher von Bedeutung. Die durch verschiedene Matrizen gefundenen niedrigen Kooperationsraten (in der Regel lt 50%) können ggf. durch diesen Quotienten erklart werden, wenn eine kompetitive Orientierung vorliegt. Mit größer werdendem P-Wert und gemeinsamer kompetitiver Orientierung wird gleichsam kompetitives Verhalten - das sich nur lohnt, wenn der Interaktionspartner sich kooperativ verhält - bestraft. Ziel von Treatmentbedingungen muß es also sein, daß Dilemma (eines hohen T-Wertes) zu relativieren. Daß - eine kooperative Interaktion anzeigende - diskriminative Hinweise vor dem Meßvorgang nicht ausreichen, deutet Tabelle 61 an. Zwar wächst mit zunehmendem Versuchungswert in den Experimenten die Eindeutigkeit einer zukunftigen kooperativen Interaktion, sie sind jedoch weitgehend durch den Versuchungswert beeinflußt. Ähnliche Hinweise liefert Tabelle 67. Effektgrößen, die durch diskriminative Hinweise vor dem Meßvorgang beeinflußt werden sollten, stehen unter der Kontrolle

des Versuchungswertes und der diskriminativen Hinweise während des Meßvorgangs. Davon weichen nur solche Studien ab, die durch Motivinduktion den Kooperationswert zu erhöhen beabsichtigen, d.h. solche Probanden werden besonders auf den Re-Wert sensibilisiert. Nur sinkt für dieses Datensample die Kooperation mit steigendem Belohnungswert für die Kooperation durch den Matrixindex. Anzunehmen ist, daß die Motivinduktion zugleich das Risiko, das mit hohen Re-Werten vorhanden ist, verdeutlicht. Hier bedarf es aber mit Sicherheit weiterer Forschungen.

Die Ergebnisse zeigen auf der Basis der Matrixwerte und den aus ihnen zu generierenden Zellenindizes kein einheitliches Bild. Die Bereitschaft zur Kooperation scheint jedoch von folgenden Bedingungen abhängig zu sein:

(a) von dem Anreiz, durch eine hohe Differenz zwischen Versuchung und Belohnung sich kompetititv zu verhalten: wenn Probanden annehmen, daß die Interaktionspartner sich auch kompetitiv verhalten oder unkontingent reagieren, dann weist ein hoher T-Wert auf einen möglichen - relativen - Gewinn durch Mogelwahlen hin

(b) Nicht-Kooperative Wahlen stehen unter der Kontrolle der P-Variblen: je kleiner der P-Wert ausfällt, desto größer ist die durch diesen Wert induzierte Strafandrohung für kompetitive Reziprozität

(c) Hohe Re-Werte führen nicht zwingend zu hohen Kooperationswerten: vermuten Probanden die Kompetitivität ihrer Ko-Akteure, dann ist der Verlust (P - Re) besonders hoch.

(ii) Anzahl der Spielzüge

Zwischen der Anzahl der Spielzüge und dem Kooperationswert konnte eine positive Beziehung aufgezeigt werden (vgl. Abbildung 27). Dieses Ergebnis liefert einen Beitrag zur Diskussion in der Literatur. BRAVER & BARNETT (1976) nehmen an, daß über viele Züge die Kooperationsbereitschaft sinkt: jeder Zug für sich genommen läßt eine kompetitive Wahl günstiger erscheinen, so daß im Verlauf der Zugfolgen die Probanden zunächst einmal eine Situation gegenseitigen Mißtrauens geschaffen haben. Um nun die Strategie zu wechseln und zu einer effektiveren, kooperativen Interaktion zu gelangen, reichen die Spielzüge nicht mehr aus (vgl hierzu auch GUMPERT, DEUTSCH & EPSTEIN, 1969. Diese Annahme wird durch die Untersuchung von WILSON (1971) in der Tendenz bestätigt. Der Autor konnte zeigen, daß auch bei kompetitiver Vorerfahrung und nachfolgenden kontingenten Kooperationstrategien die Kooperationswahlen über 300 Züge kontinuierlich anstiegen. Eine Feinanalyse zeigt, daß in den ersten 40 Zugen die Kooperationsrate in Abhängigkeit der erlebten Kompetition zunächst sank und dann anschließen steil anstieg. Die Kontrollgruppe, die aufgrund fehlender Vorerfahrung die Un-Nützlichkeit der kompetitiven Strategie nicht erleben konnte, stieg in ihrer Kooperation nur unwesentlich. Im Gegensatz zeigt eine anfängliche Kooperationsbereitschaft während der ersten Züge weniger Wirkung. Sie sank innerhalb dier Folge drastisch und stieg im Verluf der restlichen 200 Zuge

zwar an, blieb jedoch unter dem Kooperationswert der Gruppe mit kompetitiver Vorerfahrung. Die Art der Vorerfahrung ist einerseits von Bedeutung. Auf der anderen Seite gilt es, solche Bedingungen zur Verfügung zu stellen, die die Strategieänderung erleichtern. McCLINTOCK & McNEEL (1966) zeigen, daß hohe Verstärkungen und die Verfolgung der (sich einpendelnden) Reziprozität zu höheren Übergangswahrscheinlichkeiten aus einer instabilen wie aus einer kompetitiven Situation führen. Interferente Faktoren, wie nur niedrige Verstärkungsbedingungen und hohe Diskrepanzen zwischen den Gewinnen liefern niedrige Übergangwahrscheinlichkeiten.

Die mit Hilfe der Meta-Analyse zur Verfügung stehenden Daten reichen nicht aus, um die Änderung von Strategien in Abhängigkeit der durch die Zugfolge ermöglichten Erfahrungen näher zu analysieren. Stellvertretend wird hier auf die Ergebnisse von MOREHOUSE (1966) verwiesen. Er diskutiert die Parameter x = p(C(A)/C(A)C(B), y = p(C(A)/C(A)D(B), z = p(C(A)/D(A)C(B) und w = p((C(A)/D(A)D(B). Tabelle 82 gibt die Ergebnisse wieder.

Tab. 82: Mittlere Wahrscheinlichkeiten für unterschiedliche Zugfolgen aus MOREHOUSE (1966), S. 358

Übergangswahrscheinlichkeiten	Anzahl der Züge 2	5	10	300
p(C(A)/C(A)C(B)	.77	.61	.60	.73
p(C(A)/C(A)D(B)	.32	.36	.37	.50
p(C(A)/D(A)C(B)	.16	.41	.33	.24
p(C(A)/D(A)D(B)	.18	.36	.31	.20

Der Tabelle kann entnommen werden, daß in einem 2-Zug-Spiel die Wahrscheinlichkeit, auf eine CC-Erfahrung mit einer kooperativen Reaktion zu folgen, sehr hoch ist. Entsprechend fällt die Bereitschaft, hier auf eine DD-Wahl mit Kooperation zu folgen, gering aus. Überraschend ist das Verhältnis y lt z. Hier weicht das Ergbnis von den nachfolgenden Relationen ab. Demnach gilt für 2-Zug-Spiele: Wenn Teilnahmer im ersten Zug ausgebeutet wurden, zeigen sie noch zu 30% die Bereitschaft, im folgenden und letzten Zug zu kooperieren. Haben sie allerdings selbst das Spiel mit einer Mogel-Reaktion eingeleitet, bleiben sie bei dieser Strategie (p = .82).

Von Interesse ist die Entwicklung der Wahrscheinlichkeiten in Abhängigkeit kurzer (10 Züge) gegenüber langen (300 Züge) Spielen. Eine erfolgreiche Kooperationsstrategie wird zu Beginn daraufhin überprüft, ob sich der Ko-Akteur ausbeuten läßt, um ggf. zur kooperativen Strategie zurückzukehren. Die Tabelle zeigt auch, daß sich über eine lange Zugfolge Kooperationsstrategien durchsetzen lassen. Umgekehrt wird erfolgreiches Mogeln ebenso etabliert wie die Aussichtslosigkeit, auf interdependente Mogel-Wahlen mit Kooperation zu reagieren.

(iii) Spielstrategie des simultanen Mit-Spielers

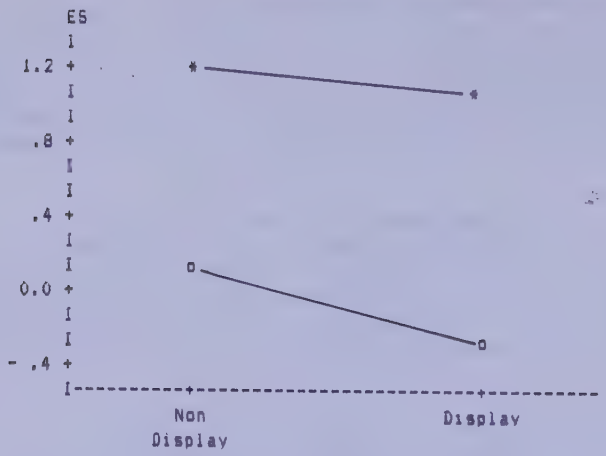
Den nicht-kontingenten Koopertionstrategien, die sich in der Höhe kooperativer Wahlen unterscheiden, stehen eine Reihe kontingenter Strategien gegenüber. Tit-for-Tat bildet die bekannteste und vordergründig einfachste Strategie: ich bin grundsätzlich zur Kooperaion bereit, wenn du jedoch mogelst, mogle ich unmittelbar und einmalig zurück. Während Abbildung 26 die Überlegenheit dieser Strategie gegenüber allen anderen Experimentalbedingungen zeigt, bildet die nachfolgende Darstellung die Effektgrößen im einzelnen ab.

Abb. 34 Stem and leaf
der Effektgrößen fur
TFT-Strategie

```
-.3 I 9
-.1 I 12 19
+.1 I 2 5 6 9 18
+.3 I 1 12
+.5 I 5 7 15
+.7 I 14
+.9 I 16
1.1 I 6 9 9 12 18 19 19
1.3 I 1 4
1.5 I 1 1
1.7 I 1 9
1.9 I 0 19
```

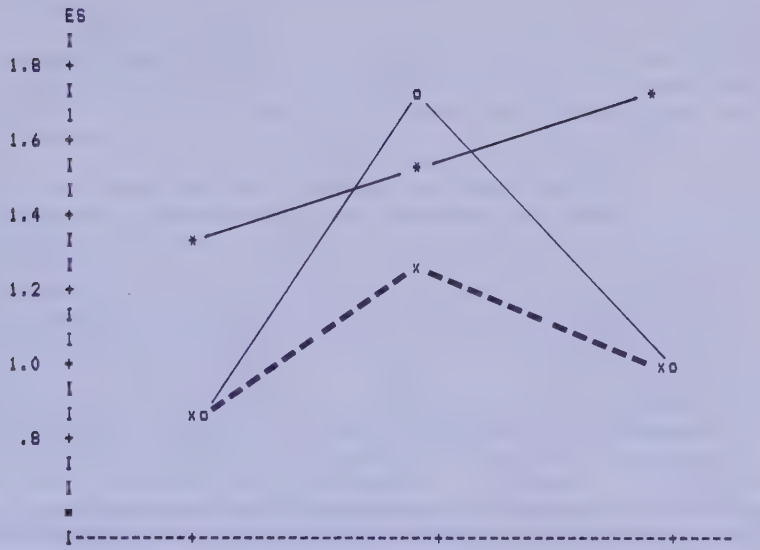
Die Abbildung zeigt eine bimodale Ausprägung, die sich durch die Effektunterschiede erklären läßt. Solche Experimente, die ohne weitere Treatments vor oder während des Meßvorgangs allein durch die TFT-Strategie des simultanen Mit-Spielers die Kooperation zu erhöhen versuchen, liefern einen mittleren Effekt von AM(ES) = 1.28. Experimente, die die TFT-Strategie anwenden, um die Wirkung anderer Variablen abzutesten, liefern demgegenüber einen Effekt von AM(ES) = .49s. Modellierung und Widerspiegelung der Punktzahl bilden hier die maßgeblichen Treatmentbedingungen. Die Ergebnisse von McNEEL (1973) zeigen den Einfluß unterschiedlicher Störfaktoren:

Abb. 40 Untersuchungsergebnisse nach McNEEL (1973)



WILSON (1971) variiert die TFT-Strategie. Als 'conciliation I' bezeichnet er eine TFT-Variante, die zwar auf Kompetition mit Kompetition antwortet, jedoch nach drei Mogelwahlen dem Interaktionspartner eine kooperative Wahl anbietet. 'Conciliation II' bietet bereits nach 2 Mogelwahlen des Interaktionspartners eine kooperative Wahl an. Übersicht 36 zeigt, daß TFT diesen Varianten überlegen ist:

Abb. 41: Kooperationseffekte in Abhängigkeit der Strategien nach WILSON (1971)



* = TFT, o = Conciliation I, x = Conciliation II

AXELROD (1980a, 1980b) bestätigt die Überlegenheit der TFT-Strategie. Er versucht in seinen Simultanexperimenten die Fragen zu beantworten, (a) unter welchen Bedingungen sich eine kooperative Strategie durchsetzen kann, (b) ob eine kooperative Strategie einer nicht-kooperativen Stregtegie überlegen ist und (c) ob sich eine kooperative Strategie auch in einem nicht-kooperativen Milieu als erfolgreich durchsetzen kann. Zu einem ersten Computer - Turnier bat er verschiedene Fachleute, Strategien zu entwickeln, die jeweils gegen sich und die anderen Strategien spielten. Von 15 am Turnier beteiligten Strategien belegten die ersten acht Plätze solche Strategien, die nicht mit Mogel-Reaktionen begannen. AXELROD bezeichnet diese Strategin als NETT. Den ersten Platz belegte Tit-for-Tat. Alle netten Strategien unterschieden sich von einander durch den Grad ihrer Vergebung infolge von Mogel-Reaktionen. TFT bildete die vergebenste Strategie: nach einer Mogelreaktion verhält sie sich kontingent, einmalig und unmittelbar durch eine Vergeltungsreaktion. Dadurch existiert bei der Interaktion mit einer TFT-Strategie auch kein 'Echo'-Effekt. Als Echo-Effekt bezeichnet AXELROD die Möglichkeit, durch die eigene Vergeltung die Mogelversuche des

Ko-Akteurs zu amplifizieren. AXELROD (1980b) teilte diese Ergebnisse der öffentlichkeit mit und bat nochmals um neue Strategien. Er erhielt 63 Einsendungen. Auch in diesem Turnier spielte je eine Strategie gegen jede und gegen sich selbst. Auch hier gewann TFT. Der Autor entdeckte jedoch durch den Vergleich der unterschiedlichen Strategie-Kombinationen ein weiteres Merkmal für den Erfolg einer Strategie: sie muß provozierbar sein, d.h. die Bereitschaft zur Vergeltung anzeigen.

E) Zum Konzept des Vertrauens und des Risikos

An früherer Stelle war bereits auf das Vertrauenkonzept verwiesen. Demnach gilt:

(1) Vertrauensinduktion besteht methodisch in der Applikation eindeutiger Hinweise zur Kooperativitat der Situation sowie von Kontingenzen, von denen angenommen werden kann, daß sie kooperative Verhaltensweisen aufrecht erhalten

(2) unkontingene, hohe Kooperationsraten wirken nicht a priori vertrauensfördernd, sie verstärken die Tendenz, sich risikobereit durch Mogelreaktionen zu verhalten

(3) Mißtrauen anzeigende diskriminative Hinweise führen nur dann zu -relativ - niedrigen Effektgrößen, wenn sie sich in den erlebbaren Kontingenzen bestätigen. Andernfalls zeigen die Bekräftigungskontingenzen den wesentlichen Einfluß.

Diese Ergebnisse sind mit den von DEUTSCH (1973) explizierten Basisannahmen des Vertrauenskonzepts identisch. Ein Ereignis hat dann positiv motivierende Bedeutung, wenn es zum Ziel beiträgt respektive Hindernisse zur Zielerreichung mindert (S. 139). Vertrauenswahlen liegen dann vor, wenn die negative Valenz größer ist als die positive und dennoch das Individuum die subjektive Wahrscheinlichkeit des Auftretens positiver Ereignisse höher veranschlägt als das Auftreten negativer Ereignisse. In Experimenten zur PDG-Forschung wird zwischen zwei Zielen unterschieden. Das 'own-gain'-orientierte Individuum legt auf einen Weg wert, soviel Gewinn wie möglich zu machen. Wenn der gunstigste Weg in der Kooperation besteht, dann kooperiert dieses Individuum. Der 'relative own gain'-orientierte Mensch sucht weniger seinen Weg nach der Gewinnmaximierung als vielmehr nach der Gewinndifferenzierung oder Differenzmaximierung. Für ihn ist die rational günstigste kooperative Strategie nicht diejenige, mit der er sein Ziel erreichen kann. Während der individuell oder kooperativ motivierte Mensch sich der Rationalität von Kosten und Nutzen stellt, entwickelt der um jeden Preis kompetitive Mensch irrationale Entscheidungen. Diesen Orientierungen ist mit zusätzlichen Bedingungen schwerer eine kooperative Wahl abzuringen. Untersuchungen, die versuchten, durch Display der bestmöglichen und real erreichten Punktzahl die Differenzmaximierung zwischen den Teilnehmern in eine Differenzminimierung zwischen möglichem und realem Punktwert zu verän-

dern, flossen in die Meta-Analyse nicht ein.

Das PDG bietet in seinem Dilemma einen ambiguenten Weg. Die negativen Valenzen durch den Versuchungswert (T) sind höher als die positiven Valenzen (Re). Die Aufgabe des Individuum besteht darin, im Verlauf des Spiels eine subjektive Wahrscheinlichkeit zu ermitteln, mit der eine bestimmte Strategie bei grundsätzlicher Bereitschaft zur kooperativen Interaktion zu einem günstigen bis maximal möglichen Ergebnis führt. Diese subjektiven Wahrscheinlichkeiten werden auf der Basis der Erfahrung entwickelt. Damit wird deutlich, daß die dikriminativen Hinweise vor der Interaktion von nur bedingter Bedeutung sind: sie weisen darauf hin, was möglich ist. Erst die Situation selbst bietet aber die ausschlaggebende Lernumwelt. Die in der Interaktion vorherrschenden Kontingenzen und Hinweise führen zur Strategie, d.h. im Sinne des Vertrauenskonzepts zur subjektiven Wahrscheinlichkeit, welche (bedingten) Reaktionen zum Ziel führen.

Ähnliches gilt für das Risiko-Verhalten. Zwar liegt die objektive Valenz im PDG beim Kompetitionswert (MINAS et al., 1960 können zeigen, daß die Mehrheit der - sich auch kooperativ verhaltenden - Probanden das PDG als kompetitiv erleben) und erst auf lange Sicht beim Kooperationswert. D.h., das PDG-Spiel ist Vertrauens- und Risiko-Spiel in einem, je nach Perspektive, aus der der Proband das Spiel betrachtet: als 'one-shot'-Interaktion, oder als Tatsache wiederholter und wiederholbarer Interaktion. Fur widerholte Interaktionen sind die kooperativen Wahlen die günstigsten und die riskohaftesten zugleich. Das Risiko kann reduziert werden, indem man von vornherein versucht, kompetitiv zu wählen. So zeigen Untersuchungen, die durch Hinweise, nicht kompetitieren zu wollen, hohe Kooperationseffekte.

Kritische Reflexion der Meta-Analyse

Mit Hilfe des von GLASS (1976) entwickelten Verfahrens sollte versucht werden, die generelle Frage zu beantworten, welche Bedingungen vor dem Hintergrund unterschiedlicher Theorien kooperatives Verhalten erklären. Damit beinhaltet die vorliegende Arbeit zwei Schwerpunkte. Zum einen sollte die Meta-Analyse, und die mit ihr verbundenen Möglichkeiten erprobt werden, zum anderen sollten mit Hilfe dieses Verfahrens solche empirische Untersuchungen systematisch und quantitativ integriert werden, die unter dem Stichwort 'Kooperation' diese interdependente Interaktionen auszuformen beabsichtigen. Insgesamt liefert die Arbeit Hinweise, daß eine meta-analytische Integration auch dann durchführbar und durchaus sinnvoll sein kann, wenn Forschungsbefunde unterschiedlichster theoretischer Provenienz systematisch zusammengefaßt werden sollen. Dennoch sind mit der Meta-Analyse verschiedene Probleme verbunden, die vor dem Hintergrund einer Anwendung aufgezeigt werden sollen. Auf der anderen Seite kann die beste Integration nur solche Evidenzen generieren, die als Informationen in den Untersuchungen vorhanden sind. Ob diese Ergebnisse, Daten und die mit ihnen verbundenen Erhebungs- und Auswertungsmethoden die 'reale Welt' kooperativer Interaktionen abbilden, ist eine zweite Frage.

A) Kritische Reflexion der Meta-Analyse

In der Diskussion zwischen GLASS (1979) und LIGHT (1979) werden zwei vermeintliche Unterschiede in der Zielsetzung einer Datenintegration hervorgehoben (vgl. auch COOK & LEVINTON, 1980): soll dieses Verfahren (primär) verwandt werden, um eine breite und summative Zusammenfassung (GLASS' Akzent) zu liefern oder bildet diese Technik eine Strategie, explorativ die Unterschiede zwischen Ergebnissen (d.h. Effektgrößen) herauszuarbeiten (LIGHT's Akzent)? Die vorliegende Arbeit zeigt, daß beide Ziele nicht voneinander zu trennen sind. Zum einen wurde aufgezeigt, welche Bedingungen Kooperation besonders wirksam bzw. weniger wirksam fördern - und im weitesten Sinne erklären. Die Analyse der Einzelstudien zielt jedoch gleichzeitig darauf ab, Unterschiede zwischen den Studien zu erfassen. Mit Hilfe multivariater Verfahren lassen sich diese Unterschiede im Hinblick auf die Unterschiede in den Effektgrößen in Beziehung setzen. Die generelle Forschungsfrage stand daher im Spannungsfeld von (fallibilistischer) Theorie- bzw. Konstruktüberprüfung und konstruktiver Elaboration von Theorien bzw. ihren Elementen, indem Interaktionen zwischen verschiedenen Merkmalen aufgedeckt werden konnten. Dies erlaubt einen theoretischen Fortschritt im Sinne "... of more precise and confident statements about segments of our world" (COOPER 1979, S. 131).

Ob diese Ergebnisse den Forschungskorpus repräsentieren, ist eine weitere Frage. Die Selektion von Studien hängt von zwei Faktoren ab: zum einen von

der Verfugbarkeit über diese Studien und zum anderen von den Selektionskriterien. Die Meta-Analyse ist bestrebt, 'alle' Studien zu einem Gegenstandsbereich zu integrieren. Damit stellt sich das triviale wie pragmatische Problem, 'alle' Studien zu eruieren. GLASS, McGAW & SMITH (1981) unterschieden hier zwischen 'Primärquellen' (Zeitschriften, Dissertationen und Veröffentlichungen unterschiedlichster Administrationen) sowie 'Sekundärquellen' (Review-Periodica, Psychological Abstract, Educational Ressources Information Center (ERIC), Dissertation Abstracts International). Insbesondere die Recherche anhand der Sekundärquellen beinhaltet organisatorische und substantielle Probleme. Trotz moderner Technologien geschieht diese Sucharbeit aufgrund beschränkter Deskriptoren weitgehend von Hand. Dies bedeutet gegebenenfalls die Lektüre Hunderter von Quellenangaben, die sehr häufig erst angefordert werden müssen. Dies gilt in der Bundesrepublik sowohl für Disserationen wie ERIC-Dokumenten. Ein anderes Such-Problem ist grundsätzlicher Art. Erfaßt werden solche Studien, die unter dem entsprechenden Distraktor aufgefunden werden. Dies schließt nicht aus, daß einund-dasselbe Phänomen unter anderen Begriffen AUCH untersucht wird. In die vorliegende Meta-Analyse flossen keine Arbeiten ein, die unter dem Begriff 'social interactions' explizit oder implizit Kooperationsphänomene untersuchten. Die Integration dieser Studien hätte vorausgesetzt, alle Literaturquellen unter diesem Suchstichwort dahin zu überprüfen, ob die beschrieben Interaktionen kooperative Interaktionen sind. GLASS, MCGAW & SMITH (1981) folgern dann auch:

"No matter how ambitious and sophisticated are one's efforts to find ALL (im Original kursiv) empirical research on a topic, the aspiration to find everything must be inevitably frustrated" (S. 63).

Diese Einschränkung impliziert, daß Schlußfolgerungen aus Meta-Analysen auch immer 'etwas tentativ' (vgl. FELDMAN 1971) bleiben werden.

Die Effektgröße wird als abhängige Variable vor dem Hintergrund der durch die Publikationen verfügbaren Informationen als unabhängige Variablen betrachtet. Die "written works of others (whether or not formally published) constitutes 'raw' data for the integrator" (FELDMAN, 1971, S. 87). Diese 'Rohdaten' können in die eigentlichen Forschungsvariablen, forschungsbezogene Ko-Variablen (oder nach GLASS, McGAW & SMITH 'substantive' Variablen) und in methodologische Variablen unterteilt werden. Mit Hilfe von Kodiersystemen wird diese Information zu quantifizieren versucht. Die einfachste Möglichkeit besteht darin, Nominalvariablen einzuführen (das Merkmal liegt vor vs. nicht vor). Häufig jedoch spiegelt diese Dichotomisierung (oder allenfalls Trichotomisierung) die unterschiedlichen Ausprägungsgrade eines Merkmals jedoch nicht wider. In der vorliegenden Arbeit wurde versucht, neben der klassifikatorischen Gegenüberstellung gemittelter Effektgrößen die Treatmentunterschiede zwischen den Experimenten acht Dimensionen (MOD, KBV, KBA, DHV, KBW, DHW, REV, PROZ-S) und zwei dichotomen Variablen (TFT, DP) zuzuordenen und auf den jeweiligen Dimensionen nach theoretischen Gesichtpunkten zu hierarchisieren. D.h., fur jede Dimension wurden Hypothesen

generiert, die die unterschiedlichen Niveaus innerhalb einer Dimension begründen. Den unterschiedlichen Niveaus wurden Rangpositionen zugeordnet und vorläufig unterstellt, die Rangabstände wären gleich. Diese Vorgehensweise erlaubte die abgestufte Betrachtung unterschiedlicher Einflußfaktoren; sie sind jedoch im Hinblick auf die unterstellten kontinuierlichen Niveauunterschiede mit Vorsicht zu interpretieren. SMITH & GLASS (1977) schlagen demgegenüber vor, mit Hilfe multidimensionaler Skalierungsmethoden Unterschiede zwischen Studien zu erfassen. So war es ihnen möglich, verschiedene Therapieformen zu zwei Superklassen zu aggregieren. In einem fünfwöchigen Seminar ordneten Experten die unterschiedlichen Therapien einander zu bzw. grenzten sie voneinander ab.

Die Effektberechnung orientierte sich in der vorliegenden Meta-Analyse am GLASS-Modell. HEDGES (1980, 1981, 1982) schlägt demgegenüber vor, die Effektgröße als Mittelwertdifferenz auf der Basis der geschätzen Populationsvarianz zu berechnen. Im Gegensatz zu der GLASS'schen (deskriptiven) Effektgröße wird hier der Effekt als Stichprobenschätzung der standardisierten Mittelwertdifferenz der Population betrachtet. Für die verzerrungsfreie mittlere Effektgröße G(u) aus einer Serie von Effektgrößen wird angenommen, daß G(u) annähernd normalverteilt ist: G(u) = N (μ, σ**2). Dies setzt jedoch eine verzerrungfreie Gewinnung der Stichprobe an Untersuchungen (oder die Grundgesamtheit) voraus. In der vorliegenden Arbeit wurde aus zwei Gründen auf die Berechnung der Effekte nach dem HEDGES-Modell verzichtet:

(i) konnte für eine große Anzahl der Untersuchungen die Populationsvarianz nicht geschätzt werden; die Bestimmung der Effektgrößen erfolgte über das von GLASS vorgeschlagene Probit-Modell

(ii) sollte nicht suggeriert werden, daß die in die Analyse eingeflossenen Studien eine definierbare Stichprobe repräsentieren. Viele Untersuchungen konnten z.B. nicht aufgenommen werden, weil eine Effektgröße nicht berechnet bzw. geschätzt werden konnte.

Diese offensichtlichen Einschränkungen lassen es insgesamt ratsam erscheinen, die vorliegenden Ergebnisse als tentative Hinweise zu betrachten, welche Konstrukte zur Erklärung von Kooperation plausibel erscheinen und welche Variablen interdependent auf Kooperation wirken.

(B) Kooperationsforschung

Der Kooperationsforschung wird vorgeworfen, sie habe sich ohne theoretischen Bezug entwickelt (vgl. weiter vorne, S. 64). Dieser Vorwurf kann zurückgewiesen werden. Alle integrierten Studien verweisen auf eine theoretische Tradition i.S. der kodierten Referenztheorien und damit auf ein mehr oder weniger explizites Netz von Hypothesen. Die Beiträge von VINACKE (1969) und McNEEL (1973) verweisen zudem auf eine Theorienkonkurrenz. Wäh-

rend VINACKE aufgrund der Forschungstradition im Rahmen der PDG-Forschung für eine feldtheoretische Integration der Forschungsbefunde plädiert, verweist McNEEL darauf, daß die Kooperationsforschung ohne intervenierende Variablen auskommt. Die vorliegende Analyse zeigt, daß die Konstrukte 'kontigente Bekräftigung' und 'diskriminative Hinweise' Kooperation besser kontrollieren als andere Konstrukte. Nicht weniger deutlich unterstützen die Ergebnisse für die TFT-Strategie diese Tatsache. Während AXELROD (1980) TFT als eine Strategie bezeichnet, die (a) nett, (b) provozierbar und (c) vergeltend reagiert, könnte man sie auch als eine Strategie bzeichnen, die (a) nie zuerst straft, (b) auf erfahrene Strafe unmittelbar, d.h. sofort und einmalig mit Strafe reagiert und (c) insgesamt kooperative Reaktionen nur bekräftigt. Sie liefert damit eine einfache und eindeutig erfaßbare Kontingenzstruktur.

Können jedoch diese Ergebnisse auf die Praxis übertragen werden? Angenommen, ein Lehrer möchte kooperative Gruppenarbeit oder eine Führungskraft kooperative Führungsstrategien durchführen bzw. anwenden, können die vorliegenden Ergebnisse herreichen, um die hohe Interdependenz zwischen Aufgabenstruktur, Interaktionsinterdependenz und den verschiedenen, die Kooperation fördernden wie hemmenden Maßnahmen so zu gestalten, daß die betroffenen Personen auch erfolgreich kooperieren? Die Antwort verlangt zunächste eine andere Antwort auf die Frage zu finden, ob in den vorliegenden Experimenten auch tatsächlich Kooperation untersucht wird oder ob andere als kooperative Verhaltensweisen gemessen werden, die fälschlicherweise als kooperative Reaktionen ausgewiesen sind. Es ist nach der Relevanz der Handlungs- und Wahlinstrumente zu fragen.

(i) Handlungsexperimente
COOK & STINGLE (1974) verweisen auf die Unterscheidung in zielorientierte Definitionen von Kooperation gegenuber synchronen motorischen Reaktionen. Abbildung 31 zeigt, daß die Mehrheit dieser Meßinstrumente in die Kategorien 'nicht-soziale bzw. soziale interdependente, alternativfreie bzw. alternativer Teilungsreaktionen' fallen. Die Abbildung weist aus, daß Personen, die ihren Ko-Akteur perzipieren können und die für alternative Reaktionen nicht belohnt werden, im Mittel eine höhere Reaktionsrate zeigen als 88.7% der Kontrollgruppe. Solche Meßinstrumente werden u.a. in den Untersuchungen von MADSEN (1967), BROTSKY & THOMAS (1967) und NACSON (1976) präsentiert. Ihnen ist gemeinsam, daß die Probanden

(a) Materialien so koordinieren müssen, daß eine definierte Aufgabe nur gemeinsam erfüllt werden kann und

(b) die Koordinationsreaktionen bei Erfüllung der Aufgabe belohnt werden.

Diese Aufgaben repräsentieren damit kooperative Interaktionen. Die Probanden müssen sowohl einen Teil ihrer instrumentellen Raktionen für die Zielerreichung zur Verfügung stellen (R(G), wie sie andereseits für die eigene Zielerreichung auf die instrumentellen Reaktionen der Mitspieler ange-

wiesen sind. Die Frage ist also weniger, ob es sich im analytischen Sinne um kooperative Interdependenzen handelt, sondern vielmehr, ob die Qualität der instrumentellen Reaktionen (Ziehen an einem Faden, Stecken von Stiften, Drücken von Tasten) die Komplexität instrumenteller Reaktionen im Alltag abbilden. Dies ist zu bezweifeln.

(ii) Relevanz der Wahlmessung

Gegenüber den Handlungsmeßverfahren zeichnen sich die Nicht-Nullsummenspiele durch einen hohen Standardisierungsgrad aus; sie wenden dieselbe Matrix-Form an, das Dilemma ist gleich struktuiert und die Unterschiede in der Präsentation des PDG können extrahiert werden. Auch wenn die Tradition der PDG-Forschung auf ein Konflikt-Paradigma zurückführbar ist, scheinen alle Experimente sich um die Frage "What has an effect on the frequency of the C choice and to what extent?" (RAPOPORT 1967, S. 462) zu zentrieren. Dennoch bestreitet RAPOPORT selbst die Möglichkeit, Ergebnisse aus der Laborforschung auf reale Situationen zu extrapolieren. Er führt dies u.a. auf die eingeschränkten (konvergenten) Designs zurück, in denen eine abhängige Variable - und dies in der Regel die gemittelte und konfundierte Wahrscheinlichkeit kooperativer Wahlen - in Abhängigkeit verschiedener unabhängier Variablen untersucht wird. Statt dessen sollten mit Hilfe divergenter Designs eine Anzahl unabhängiger Variablen konstant gehalten werden, um ihren Einfluß auf eine große Anzahl abhängiger Variablen zu kontrollieren. In die vorliegende Meta-Analyse wurden die in den Untersuchungen dargestellten Prozentwerte kooperativer Reaktionen

$$p(C) = \frac{C(A) + C(B)}{2 * Z} \qquad (Z = \text{Anzahl der Züge})$$

integriert. Die auf dieser Basis ermöglichten Aussagen sind äußerst eingeschränkt, da sie weder über bedingte Reaktionen noch über Lernfortschritte während des Spiels informieren. Ebenfalls können keine Aussagen darüber gemacht werden, welche (externen) Bedingungen Einfluß auf bedingte Reaktionen und Lernprozesse ausüben. Hierzu bedarf es anderer abhängiger Variablen, die zwar spätestens seit RAPOPORT (1967) zur Diskussion stehen, aber in den Untersuchungen kaum zur Anwendung kommen:

(a) der Gewinn des einen Partners in Relation zum Gewinn des anderen Akteurs; diese Variable verwendet AXELROD (1980a,b) in seinen Turnieren

(b) bedingte Reaktionsfrequenzen (C/CC; C/DC; C/CD; C/DD)

(c) Wahrscheinlichkeiten, mit denen Reaktionen geandert bzw. beibehalten werden:

D/CC	vs	C/CC
D/CD	vs	C/CD
C/DC	vs	D/DC
C/DD	vs	D/DD

(d) die Analyse von iterativen Spielverläufen mit Hilfe von Markov-Ketten; dies erlaubt die Analyse von Strategieänderungen und/oder Lernprozessen.

Wenn auch die vorliegenden Kooperationskonstrukte, gemessen an Real-Situation und möglichen operationalen Verwirklichungen, als invalide bezeichnet werden können, so liefern sie dennoch Hinweise über die Bedingungen, unter denen Personen sich kooperativ verhalten. Nur sind diese Aussagen sehr allgemein und in ihrem Bezug zur Realität nicht überprüft. Eine weitere Einschränkung, die diese Meta-Analyse aufgrund der Untersuchungen unterliegt, besteht darin, ein 'black-box-Modell' als kleinsten gemeinsamen Nenner, mit dem die Einzelstudien verglichen respektive unterschieden werden konnten, unterstellt zu haben. Zukünftige Kooperationsforschung sollte sich eines weniger sparsamen Modells bedienen, da kooperative Interaktionen nicht nur durch externale Bedingungen, sondern auch durch die Verarbeitung dieser Bedingungen gesteuert wird. Die Kooperationsforschung liefert hierzu jedoch kaum Beiträge.

Literaturverzeichnis

ADCOCk, C.J. 1960
A note on combining probabilities
Psychometrika, 25, 303 - 305

AUSUBEL, D.P. & THOL, N.J. 1983
Organisationshilfen als wirksame Lernstrategien?
Psychologie in Erziehung und Unterricht, 30, 64 - 66

AXELROD, R. 1980a
Effective choice in the Prisoner's Dilemma
Journal of Conflict Resolution, 24, 3 - 25

AXELROD, R. 1980b
More effective choice in the Prisoner's Dilemma
Journal of Conflict Resolution, 24, 379 - 403

AXELROD, R. & HAMILTON, W.D. 1981
The evolution of cooperation.
Science, 211, 1390 - 1396

AZRIN, N.H. & LINDSLEY, O.R. 1956
The reinforcement of cooperation between children
Journal of Abnormal and Social Psychology, 52, 100 - 102

BAKAN, D. 1966
The test of significance in psychological research.
Psychological Bulletin, 66, 423 - 437

BANDURA, A. 1979
Aggression. Eine sozial-lerntheoretische Analyse
Stuttgart;
original: Aggresion: A social learning analysis. 1973
New Jersey

BANDURA, A. 1979
Sozial-kognitive Lerntheorie.
Stuttgart;
original: Social learning theory 1977
New Jersey

BANGERT, R.L., KULIK, J.A., & KULIK, C.-L.C. 1983
Individualized system of instruction in secondary schools.
Review of Educational Research, 53, 143 - 158

BARNES, B.R. & CLAWSON, E.V. 1975
Do advance organizers facilitate learning?
Review of Educational Research, 45, 637 - 659

BASS, B.A. 1976
Aggressive andprosocial television. Programs and their effects on cooperation and competition in young children.
Dissertation: University of Michigan

BERGMANN, G. 1957
Philosophy of science.
Madison

BERKE, R.E. 1977
The relationship between co-operation and competition in social life: A general system theory analysis.
Dissertation: University of Southern California

BIRNBAUM, A. 1954
Combining independent tests of significance.
American Statistical Association Journal, 49, 559 - 574

BIXENSTINE, V.E., POTASH, H.M., & WILSON, K.V. 1963
Effects of level of cooperation choice by the other player on choices in a Prisoner's Dilemma Game. Part I
Journal of Abnormal and Social Psychology, 66, 308 - 313

BORG, G.R. & GALL, M.D. 1979
Educational Research: An introduction.
New York

BOZARTH, J.D., ROBERTS, Jr., R.R. 1972
Signifying significant significance.
American Psychologist, 27, 774 - 775

BRAVER, S.L. & BARNETT, B. 1976
Effects of modeling on cooperation in a prisoner's dilemma game.
Journal of Personality and Social Psychology, 33, 161 - 169

BRAVER, S.L. & VAN ROHRER 1978
Superiority of vicarious over direct experience in interpersonal conflict resolution.
Journal of Conflict Resolution, 22, 143 - 155

BRION-MEISEL, S.J. 1977
Helping, sharing, and cooperation: An intervention study of middle childhood.
Dissertation: University of Utah

BROTSKY, S.J. & THOMAS, K. 1967
Cooperative behavior in preschool children.
Psychonomic Science, 9, 337 - 338

BROZEK, J. & TIEDE, K. 1952
Reliable and questionable significance in a series of statistical tests.
Psychological Bulletin, 49, 339 - 341

BRYAN, J.H. 1975
Children's cooperation and helping behaviors.
In: HETHERINGTON, E.M. (Ed.), Review of child development research, Vol. 5, 127 - 181, Chicago

BRYAN, J.H. & LONDON, P. 1970
Altruistic behavior by children.
Psychological Bulletin, 73, 200 - 211

BRYAN, J.H. & WALBECK, N.H. 1970a
The impact of words and deeds concerning altruism upon children.
Child Development, 41, 747 - 757

BRYAN, J.H. & WALBECK, N.H. 1970b
Preaching and practicing generosity: Children's actions and reactions.
Child Development, 41, 329 - 353

BUELL, J., STODDARD, P., & HARRIS, F.R. 1968
Collateral social development accompanying reinforcement of outdoor play in a preschool child.
Journal of Applied Behavior Analysis, 1, 167 - 173

CAMPBELL, D.T. 1972
On the genetics of altruism and the counter-hedonic components in human culture.
Journal of Social Issues, 28, 21 - 37

CAMPBELL, D.T. & STANLEY, J.C. 1963
Experimental and quasi-experimental designs for research on teaching.
In: GAGE, N.L. (Ed.), Handbook of research on teaching.
Chicago

CARLTON, M. et al. 1975
Die Auswirkung von Szenen zum sozialen Lernen aus der Fernsehreihe "Sesamstraße" auf Vorstellungsinhalte und Spielverhalten von Kindern.
Zeitschrift für Sozialpsychologie, 6, 348 - 358

CARVER, D.P. 1978
The case against statistical significance testing.
Harvard Educational Review, 48, 378 - 399

COHEN, D.J. 1962
Justin in his peers: An experimental analysis of a child s social work.
Child Development, 23, 697 - 717

COHEN, J. 1966
Statistical power analysis for behavioral science.
New York

COHEN, P.A. 1981
Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies.
Review of Educational Research, 51, 281 - 309

COHEN, P.A., EBERLING, B.J., & KULIK, J.A. 1981
A meta-analysis of outcome studies of visual-based instruction.
Educational Communication and technology, 29, 26 - 36

COOK, H. & STINGLE, S. 1974
Cooperative behavior in children.
Psychological Bulletin, 81, 918 - 933

COOK, T.D. & CAMPBELL, D.T. 1979
Quasi experimentation. Design and analysis issues for field settings.
Chicago

COOK, T.D. & LEVITON, L.C. 1980
Reviewing the literature: A comparison of traditional methods with meta-analysis.
Journal of Personality, 48, 449 - 472

COOPER, L. et al. 1980
The effects of cooperative, competitive, and individualistic experiences on interpersonal attraction among heterogenous peers.
Journal of Social Psychology, 111, 243 - 252

COOPER, H.M. 1979
Statistically combining independent studies: A meta-analysis of sex differences in conformity research.
Journal of Personality and Social Psychology, 37, 131 - 146

COOPER, H.M. & ROSENTHAL, R. 1980
Statistical versus traditional procedures for summarizing research findings.
Psychological Bulletin, 87, 442 - 449

CRONBACH, L.J. & SNOW, R.E. 1977
Aptitudes and instructional methods.
New York

CROTT, H.W. 1971
Experimentelle Untersuchungen zum Verhandlungsverhalten in kooperativen Spielen.
Zeitschrift für Sozialpsychologie, 2, 61 - 74

CROTT, H.W. 1972
Der Einfluß struktureller und situativer Merkmale auf das Verhalten in Verhandlungssituationen.
Zeitschrift für Sozialpsychologie, 3, 134 - 158

CROTT, H.W. 1979
Soziale Interaktionen und Gruppennormen.
Stuttgart

DAHRENDORF, R. 1958
Homo sociologicus. I.Allgemeiner Teil.
Kölner Zeitschrift für Soziologie und Sozialpsychologie, 10, 178 - 208

DANIEL, W.J. 1942
Cooperative problemsolving in rats.
Journal of Comparative Psychology, 34, 361 - 368

DEUTSCH, M. 1949
A theory of co-operation and competition.
Human Relations, 2, 129 - 152

DEUTSCH, M. 1958
Trust and suspiction.
Journal of Conflict Resolution, 2, 265 - 279

DEUTSCH, M. 1962
Cooperation and trust: Some theoretical notes.
In: JONES, M.R. (Ed.), Nebraska Symposion on Motivation. 275 - 297.
Lincoln

DEUTSCH, M. 1965
The effects of cooperation and competition upon group process.
In: POSHANSKY, H. & SEIDENBERG, B. (Eds.), Basic studies in social psychology. 552 - 577.
New York

DEUTSCH, M. 1968
Field theory in Social Psychology.
In: LINDZEY, G. & ARONSON, E. (Eds.), The handbook of social psychology, Vol. I, 412 - 487

DEUTSCH, M. 1973
The resolution of conflict.
Yale
deutsch: Konfliktregulierung: Konstruktive und destruktive (1976)
Prozesse.
München

DIXON, W.J. & MASSEY, F.J. 1969
Introduction to statistical analysis.
New York

DRAPER, N.R. & SMITH, H. 1966
Applied regression analysis.
New York

DUBIN, R. & TAVEGGIA, T.C. 1968
The teaching-learning paradox.
Oregon

DULANY, D.E. 1968
Awareness, rules, and propositional control: A confrontation with S-R-behavior theory.
In: DIXON, T.R. & HORTON, D.L. (eds.), Verbal behavior and general behavior theory.
New Jersey

DUNKIN, M.J. & BIDDLE, B.J. 1974
The study of teaching.
New York

DURKHEIM, E. 1887
Le suicide.
deutsch: Der Selbstmord (1973)

DURKHEIM, E. 1893
De la division du travail social.
Paris
deutsch: Über die Teilung der sozialen Arbeit (1977)
Frankfurt a.M.

DURKHEIM, E. 1895
Les regles de la method sociologique.
Paris
deutsch: Regeln der soziologischen Methode (1961)

DURKHEIM, E. 1912
Le formes elementaires de la vie religieuse.
Paris
deutsch: Die elementaren Formen des religiösen Lebens. (1981)
Frankfurt a.M.

EDGINGTON, E.S. 1972
An additive method for combining probability values from independent experiments.
The Journal of Psychology, 80, 351 - 363

EDGINGTON, E.S. 1972b
A normal curve method for combining probability values from independent experiments.
The Journal of Psychology, 82, 85 - 89

EIFERMANN, R. 1970
Cooperativeness and egalitarism in kibbutz children's game.
Human Relations, 23, 579 - 587

ELASHOFF, J.D. & SNOW, R.E. 1971
Pygmalion reconsider - A case study in statistical inference: Reconsideration of the Rosenthal-Jacobson data on teaching expactancy.
Belmont, California

EYSENCK, H.J. 1978
An exercise in mega-silliness.
American Psychologist, 33, 517

FARQUHAR, W.W. & KRUMBOLTZ, J.D. 1959
A check list for evaluating experimental research in psychology and education.
Journal of Educational Research, 52, 353 - 354

FEDER, J.L. 1978
A development study of the effects of external and self reinforcement of cooperative models on the social interaction of dyads.
Dissertation: University of Denver

FELDMAN, K.A. 1971
Using the work of others: Some observations on reviewing and integration.
Sociology of Education, 44, 86 - 102

FEDMAN, K.A. & NEWCOMB, T.M. 1970
The impact of college on students.
San Francisco

FEYERABEND, P. 1975
Wider den Methodenzwang. Skizze einer anarchistischen Erkenntnistheorie.
Frankfurt a.M.

FISHER, R.A. 1948
Combining independent tests of significance.
The American Statistican, 2, 30

FLANDERS, J.P. 1968
A review of research on imitativ behavior.
Psycholigical Bulletin, 69, 316 - 337

FRICKE, R. 1977
Über die Bedeutung von Teststärkenanalysen in der empirischen Unterrichtsforschung.
Zeitschrift für empirische Pädagogik, 1, 2 - 27

FRICKE, R. 1978
Möglichkeiten zur zusammenfassenden Darstellung von unabhängigen Forschungsergebnissen zur Lehrer-Schüler-Interaktion.
Zeitschrift für erziehungswissenschaftliche Forschung, 11, 208 - 215

GAGE, N.L. 1978
The scientific basis of the art of teaching.
New York
deutsch: Unterrichten: Kunst oder Wissenschaft (1979)
München

GALLO Jr., P.S. 1976
Meta-analysis - A mixed meta-phor?
American Psychologist, 3, 515 - 517

GALLO Jr., P.S. & WINCHELL, J.D. 1970
Matrix indices, large rewards, and cooperative behavior in a prisoner's dilemma game.
Journal of Social Psychology, 81, 235 - 241

GARNER, K. & DEUTSCH, M. 1974
Cooperative behavior in dyads.
Journal of Conflict Resolution, 18, 634 - 645

GEWIRTZ, J.L. 1969
Mechanism of social learning: some roles of stimulation and behavior in early human development.
In: GOSLIN (Ed.), Handbook of socialization theory and research, Chicago

GEWIRTZ; J.L. & BEAR, D.M. 1958
The effects of brief social deprivation on behaviors for a social reinforcer.
Journal of Abnormal and Social Psychology, 56, 49 - 56

GIESEN, B. & SCHMID, M. 1978
Methodologische Modelle und soziologische Theorien.
In: HONDRICH, O. & MATTHES, J. (Hrsg.), Theorienvergleich in den Sozialwissenschaften, 232 - 254.
Darmstadt, Neuwied

GLASS, G.V: 1976
Primary, seconday, and meta-analysis of research.
Educational Researcher, 10, 3 - 8

GLASS, G.V. 1978a
Reply to MANSFIELD & BUSSE.
Educational Researcher, 7, 3

GLASS, G.V. 1978b
Integrating findings: the meta-analysis of research.
IN: SHULMAN, L. (Ed.), Review of Research in Education, 5, 351 - 377

GLASS, G.V. 1979
Policy for the unpredictable. Uncertainty research and policy.
Educational Reseacher, 8, 12 - 14

GLASS, G.V. 1980
On criticism of our class size/student achievement research: No points concerned.
Phi Delta Kappan, 62, 242 - 244

GLASS, G.V. & SMITH, M.L. 1977
Meta-analysis of psychotherapy outcome studies.
American Psychologist, 32, 752 - 760

GLASS, G.V. & SMITH, M.L. 1978
Reply to Eysenck.
American Psychologist, 33, 517 - 519

GLASS, G.V. & SMITH, M.L. 1979
Meta-analysis of research on the relationship of class size and achievement.
Educational Evaluation and Policy Analysis, 1, 2 - 16

GLASS, G.V., McGAW, B., & SMITH, M.L. 1981
Meta-analysis in social science.
Beverly Hills

GOULDNER, A.W. 1960
The norm of recpiprocity: A preliminary statement.
American Sociological Review, 25, 161 - 178

GRECA, R. 1978
Die makrologische Rekonstruktion sozialwissenschaftlicher Theorien am Beispiel von Forschungsarbeiten zur "Wahrnehmung sozialer Ungleichheit" - Ein empirisches Verfahren des Theorienvergleichs.
In: HONDRICH, O. & MATTHES, J. (Hrsg.), Theorienvergleich in den Sozialwissenschaften, 255 - 284
Darmstadt und Neuwied

GREENWALD, A.G. 1975
Consequences of prejudice against the null hypothesis.
Psychological Bulletin, 82, 1 - 20

GROEBEN, N. & WESTMEYER, H. 1975
Kriterien psychologischer Forschung.
München

GUMPERT, P., DEUTSCH, M., & EPSTEIN, Y. 1969
Effect of incentive magnitude on cooperation in the prisoner's dilemma game.
Journal of pertsonality and Social Psychology, 11, 66 - 69

HAERTEL, G.D. & WALBERG, H.J. 1980
Investigating an educational productivity model.
Evaluation in Education, 4, 103 - 104

HAERTEL, G.D., WALBERG, H.J., & HAERTEL, E.H. 1980
Classroom socio-psychological environment.
Evaluation in Education, 4, 113 - 114

HALL, J.A. 1978
Gender effects in decoding nonverbal cues.
Psycholigical Bulletin, 85, 845 - 857

HALL, J.A. 1980
Gender differences in skill and sending and integrating non-verbal emotional cues.
Evaluation in Education, 4, 71 - 72

HAKE, D.F., POWELL, J., & OLSEN, R. 1969
Conditioned surpression as a sensitive baseline for social facilitation.
Journal of experimental Analysis of Behavior, 12, 807 - 816

HAKE, D.F. & VUKELICH, R. 1972
A classification and review of cooperation procedures.
Journal of the Experimental Analysis of Behavior, 18, 333 - 343

HAKE, D.F., VUKELICH, R. & KAPLAN, S.J. 1973
Audit reponses: Responses maintained by access to existing self or coaktor scores during non-social, parallel work, and cooperation procedures.
Journal of the Experimental Analysis of Behavior, 19, 409 - 423

HAKE, D.F., VUKELICH, R., & OLVERA, D. 1975
The measurement of sharing and cooperation as equity effects and some relationsip between them.
Journal of Experimental Analysis of Behavior, 23, 63 - 79

HAKE, D.F. & OLVERA, D. 1978
Cooperation and competition and related phenomena.
In: CATANIA, A.C. & BRIGHAM, T.A. (Eds.), Handbook of applied behavior analysis, New York

HANSON, N.R. 1965
Pattern of discovery.
Cambridge

HARRISON, A.A. & McCLINTOCK, C.G. 1965
Previous experience within the dyad and cooperative behavior.
Journal of Personality and Social Psychology, 1, 671 - 675

HARTUP, W.W. 1965
Social behavior of children.
Review of Educational Research, 35, 122 - 129

HARTUP, W.W. 1970
Peer interaction and social organization.
In: MUSSEN, P.H. (Ed.), Carmichael's anual of child psychology,
Vol. 2, New York

HECKHAUSEN, H. 1980
Motivation und Handeln
Berlin, Heidelberg, New York

HEDGES, L.V. 1978
Estimates of effect size based on sample standardization
of a mean difference.
Manuskript

HEDGES, L.V. 1979
A meta-analysis of studies of four teacher behavior variables.
Standford

HEDGES, L.V. 1981
Distribution theory for GLASS's estimator of effect size and
related estimators.
Journal of Educational Statistics, 6, 107 - 128

HEDGES, L.V. 1982
Fitting categorial models to effect size from a series of
experiments.
Journal of Educational Statistics, 7, 119 - 137

HEDGES, L.V. 1982
Fitting continious models to effect size data.
Journal of Educational Statistics, 7, 245 -270

HEDGES, L.V. 1982
Estimation of effect size from a series of indenpendent
experiments.
Psychological Bulletin, 92, 490 - 499

HEDGES, L.V. 1983
A random effects model for effect sizes.
Psychological Bulletin, 93, 388 - 395

HEDGES, L.V: & OLKINS, I. 1979
Three vote counting methods for the estimation of effect size and statistical significance of combined results.
AREA-Papier

HEDGES, L.V. & OLKIN, I. 1980
Vote counting methods in research synthesis.
Psychological Bulletin, 88, 359 - 369

HEDGES, L.V. & OLKIN, I. 1983
Clustering estimates of effect magnitude from independent studies.
Psychological Bulletin, 93, 563 - 573

HEDGES, L.V. & STOCK, W. 1983
The effect of class size: An examination of rival hypotheses.
American Educational Research Journal, 20, 63 - 85

HILGARD 1978
Open peer commentary to ROSENTHAL & RUBIN: Interpersonal expectancy effects: the first 345 studies.
Bhavioral and Brain Science, 3, 421

HINDE, C. 1974
Biological bases of human social behavior.
New York

HOFSTADTER, D.R. 1983
Metamagicum: Kann sich in einer Welt voller Egoisten kooperatives Verhalten entwickeln?
Spektrum der Wissenschaft, Heft 8, 8 - 14

HOECKELE, C.R. 1972
The development of cooperative and competitive behavior in children.
Dissertation: Hafstra University

HOGAN, J.L., FISHER, R.H., & MORRISON, B.J. 1974
Social feedback and cooperative game behavior
Psychological Reports, 34, 1075 - 1082

HOLLAND, J.G. & SKINNER, B.F. 1961
The analysis of behavior.
New York
deutsch: Analyse des Verhaltens (1974)
Munchen

HOMANS, G.C. 1961
Social behavior. Its elementary forms.
New York

HORNKE, L.F. 1973
Verfahren zur Mittelung von Korrelationen.
Psychologische Beiträge, 15, 87 - 105

HORNKE, L.F. 1981
Erwerb prosozialen Verh/ltens durch Fernsehsendungen:
Gültigkeits- und Integrationsprobleme empirischer Befunde zur
Vorschulserie Sesamestraße.
Unveröffentlichtes Manuskript, Universität Düsseldorf

HORNKE, L.F. 1983
Integration empirischer Forschungsergebnisse?
Psychologie in Erziehung und Unterricht, 30, 54 - 63

HORNSTEIN, H.A. 1972
Promotive tension: The basis of prosocial behavior from
a Lewinian perspective.
Journal of Social Issues, 28, 191 - 218

HUNT, K. 1975
Do we really need more replications?
Psychological Reports, 36, 587 - 593

HUNTER, J.E. & SCHMIDT, F.L. 1978
Differential and single group validity of employment tests
by race: A critical analysis of three recent studies.
Journal of Applied Psychology, 63, 1 -11

HUNTER, J.E., SCHMIDT, F.L. & HUNTER, R. 1979
Differential validity of employment tests by race: A
comprehensive review and analysis.
Psychological Bulletin, 86, 721 - 735

IDE, J.C. et al. 1980
Peer influences.
Evaluation in Education, 4, 111 - 112

IGELMO, C.J. 1976
The effects of peer model's age and sharing behavior on preschool boys' cooperation in a conflict situation.
Dissertation: University of Kentucky

INGENKAMP, K.H. 1970
Handbuch der Unterrichtsforschung. Teil 1
(Deutsche Bearbeitung von GAGE, N.L. (ed.), Handbook of rsearch on teaching, chapter 1 - 9)
Weinheim

IRLE, M. 1975
Lehrbuch der Sozialpsychologie.
Göttingen

IVERSON, B.K. & WALBERG, H.J. 1980
Home environment.
Evaluation in Education, 4, 107 - 108

IVERSON, B.K. & WALBERG, H.J. 1979
Home environment and learning: A quantitative synthesis.
ED 168 714

JACKSON, G.B. 1980
Methods for integrative reviews.
Review of Educational Research, 50, 438 - 460

JEHLE, P. 1978
Trainingskurs: Verhaltenstheorie. Band 1
Düsseldorf

JENSEN, R.E. & MOORE, S.G. 1977
The effect of attribute statements on cooperativeness and competitiveness in school-age boys.
Child Development, 48, 305 - 307

JERSILD, A.T. 1960
Child Psychology.
New Jersey

JOHNSON, D.W. & JOHNSON, R.T. 1974
Instructional goal structure: Cooperative, competitive, or individualistic.
Review of Educational Research, 44, 213 - 240

JOHNSON, D.W. & JOHNSON, R.T. 1975
Learning together and alone: Cooperation, competition, and individualization.
New York

JOHNSON, D.W. et al. 1981
Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis.
Psychological Bulletin, 89, 47 - 62

JOHNSON Jr., G.B. 1957
A method for evaluating research articles in education.
Journal of Educational Researcher, 51, 149 - 151

JONES, L.V. & FISKE, D.W. 1953
Models for testing the significance of combined results.
Psychological Bulletin, 50, 375 - 382

KAGAN, S. & MADSEN, M.C. 1971
Cooperation and competition of Mexican, Mexican-American, and Anglo-American children of two ages under four instructional sets.
Developmental Psychology, 5, 32 - 39

KANFER, F.H. & PHILLIPS, J.S. 1970
Learning foundations of behavior therapy.
New York

KATZ, L.G. 1968
A study of the changes in behavior of children enrolled in two types of Head Start classes.
Dissertation: Standford University

KAUFMANN, H. 1970
Aggression and altruism.
New York

KAUFMANN, A., BARON, A. & KOPP, R.E. 1966
Some effects of instruction on human operant behavior.
Psychonomic Monographs Supplements, 1, 243 - 250

KAZDIN, A.E. & ROGERS, T. 1978
On paradigms and recycled ideology: Analogue research revisited.
Cognitive Therapy and Research, 2, 105 - 117

KELLEY, H.H. et al. 1962
The development of cooperation in the 'minimal social situation'.
Psychological Monographs, 76, Whole-No. 538

KELLEY, H.H. & STAHELSKI, A.J. 1970
The social interaction basis for cooperators and competititors beliefs about others.
Journal of Personality and Social Psychology, 16, 66 - 91

KERLINGER, F.N. 1964
Foundations of behavioral research.
New York

KERLINGER, F.N. & PEDHAZUR, E.J. 1973
Multiple regression in behavioral science.
New York

KIRBY, F.D. & TOLER Jr., H.C. 1970
Modification of preschool isolate behavior: A case study.
Journal of Applied Behavior Analysis, 3, 309 - 314

KLAUER, K.J. 1981
Zielorientiertes Lehren und Lernen bei Lehrtexten. Eine Meta-Analyse.
Unterrichtswissenschaft, 4, 300 - 318

KLIPPSTEIN, E. & KLIPPSTEIN, H. 1978
Soziale Erziehung mit kooperativen Spielen: Regelspiele für Vorschule und Schule.
Bad Heilbronn

KOMORITA, S.S. 1978
Evaluating coalition theories.
Journal of Conflict Resolution, 22, 691 - 706

KREMER, B.K. et al. 1980
Science evaluation research.
Evaluation in Education, 4, 125 - 129

KRIVOHLAVY, J. 1974
Zwichenmenschliche Konflikte und experimentelle Spiele.
Bern, Stuttgart, Wien

KUHN, T.S. 1969
Die Struktur wissenschaftlicher Revolutionen.
Frankfurt a. M.

KULIK, J.A., KULIK, C.-L.C., & COHEN, P.A. 1979
Research on audio tutorial instruction: A meta-analysis of cooperative studies.
Research in Higher Education, 11, 321 - 341

KULIK, J.A., KULIK, C.-L.C., & COHEN, P.A. 1979
A meta-analysis of outcome studies of Keller's personalized system of instruction.
American Psychologist, 34, 307 - 318

KULIK, J.A., COHEN, P.A., & EBERLING, B.J. 1980
Effectiveness of programmed instruction in higher education: A meta-analysis of findings.
Educational Evaluation and Policy Analysis, 2, 51 - 64

KULIK, J.A. & KULIK, C.-L.C. 1980
Individualized college teaching.
Evaluation in Education, 4, 64 - 67

KULIK, J.A., KULIK, C.-L.C. & COHEN, P.A. 1980
Effectiveness of computer based college teaching: A meta-analysis of findings.
Review of Educational Research, 50, 525 - 544

LADAS, H. 1980
Summarizing research: A case study.
Review of Educational Research, 50, 597 - 624

LAKATOS, I. 1974
Falsifikation und die Methodologie wissenschaftlicher Forschungsprogramme.
In: LAKATOS, I. & MUSGRAVE, A. (Hrsg.), Kritik und Erkenntnisforschritt.
Braunschweig

LENG, R.J. & WALKER, S.G. 1982
Comparing two studies of crisis bargaining.
Journal of Conflict Resolution, 26, 571 - 591

LEWIN, K. 1939
Field theory and experiment in social psychology.
American Journal of Sociology, 44, 868 - 897

LIGHT, R.J. 1979
Capitalizing on variation: How conflicting research findings can be helpful for policy.
Educational Researcher, 8, 7 -11

LIGHT, R.J. & SMITH, P.V. 1971
Accumulating evidence: Procedures for resolving contradicitons among different research studies.
Havard Educational Review, 41, 429 - 471

LIGHT, R.J. & PILLEMER, D.B. 1982
Numbers and narrative: Combining their strengt in research reviews.
Havard Educational Review, 52, 1 - 26

LINDSKOLD, S., CULLEN, P., GAHAGAN, J., & TEDESCHI, J.T. 1970
Experimental aspects of reaction to positive inducements.
Developmental Psychology, 3, 277 - 284

LINDSKOLD,S. & COLLINS, M.G. 1978
Inducing cooperation by groups and individuals.
Journal of Conflict Resolution, 22, 679 - 690

LINDSLEY, O.R. 1966
Experimental analysis of cooperation and competition.
In: VERHAVE, T. (Ed.), The experimental analysis of behavior.
New York

LOEB, H.J.W. 1975
Social interactions and performance under competitive and cooperative working conditions: A developmental study of elementary school children.
Dissertation: Bryan Mawr College

LUBORSKY, L. et al. 1971
Factors influencing to outcome of psychotherapy: A review of quantitative research.
Psychological Bulletin, 75, 145 - 185

LUCE, R.D. & RAIFFA, H. 1957
Games and decision.
New York

LUITEN, J. et al. 1979
The advance organizer: A review of research using GLASS's technique of meta-analysis.
ED 171 803

LUITEN, J., AMES, W., & ACKERSON, G. 1980
A meta-analysis of the effect of advance organizers on learning and retention.
American Educational Research Journal, 17, 211 - 218

LYKKEN, D.T. 1968
Statistical significance in psychological research.
Psychological Bulletin, 70, 151 - 159

LYSAKOWSKI, R.S. & WALBERG, H.J. 1980
Classroom reinforcement.
Evaluation in Education, 4, 115 - 116

MACH, E. 1911
Die Analyse der Empfindungen.
Jena

MADSEN, M.C. 1967
Cooperative and competitive motivation of children
in three mexican sub-cultures.
Psychological Reports, 20, 1307 - 1320

MAHONEY, M.J. 1974
Cognition and behavior modification.
Cambridge
deutsch: Kognitive Verhaltentherapie (1977)
München

MANSFIELD, R.S. & BUSSE, T.V. 1977
Meta-anlysis of research: A rejoinder to GLASS.
Educational Researcher, 6, 3

MARTIN, R.J. 1974
Teaching children to cooperate: An experimental study
using non-zero-sum games.
Dissertation: De Paul University

MARWELL, G., SCHMITT, D.R. & SHOTOLA, R. 1971
Cooperation and interpersonal risk.
Journal of Personality and Social Psychology, 18, 9 - 32

McCLINTOCK, C.G. & McNEEL, S.P. 1967
Prior dyadic experience and monetary reward as determinants
of cooperative and competitive game behavior.
Journal of Personality and Social Psychology, 5, 282 - 294

McCLINTOCK, C.G. & McNEEL, S.T. 1966
Reward and score feedback as determinants of cooperative
and competitive game behavior.
Journal of Personality and Social Psychology, 6, 606 - 613

McCLINTOCK, C.G., MOSKOWITZ, J.M., & McCLINTOCK, E. 1977
Variation in preferences for individualistic, competitive,
and cooperative outcomes as a function of age, game class,
and task in nursery school children.
Child Development, 48, 1080 - 1085

MCGAW, B. & GLASS, G.V. 1980
Choice of the metric for effect size in meta-analysis.
American Educational Research Journal, 17, 325 - 337

McGLYNN, R.P. 1982
A comment on the meta-analysis of goal structure.
Psychological Review, 92, 184 - 185

McNEEL, S.T. 1973
Training cooperation in the prisoner's dilemma.
Journal of Expectional Social Psychology, 9, 335 - 348

MEAD, M. 1937
Cooperation and competition among primitive people.
New York

MINAS, J.S., SCODEl, A., & MARLOWE, D. 1960
Some descriptive aspects of two-person non-zero-sum games.
Journal of Conflict Resolution, 4, 193 - 198

MISCHEL, W. 1968
Personality and assessment.
New York

MITHAUG, D.: 1969
The development of cooperation in alternative task situations
Journal of Experimental Child Psychology, 8, 443 - 460

MOELLER, T.P. 1974
Cooperative behaviors of four-year-old girls in nursery school settings.
Dissertation: University of Michigan

MOREHOUSE, L.G. 1966
One-play, two-play, and ten-play runs in PDG.
Journal of Conflict Resolution, 10, 354 - 362

MOSTELLER, F. & BUSH, R.R. 1954
Selected quantitative techniques.
In: LINDZEY, G. (Ed.), Handbook of Social Science, 289 - 334
Cambridge

NACSON, L.R. 1976
The effects of contingencies of reinforcement, instructional set and player sex on cooperative and competitive game behavior.
Dissertation: Columbia University

NELSON, L. & MADSEN, M.C. 1969
Cooperation and competition in four-year-olds as a function of reward contingency and subculture.
Developmental Psychology, 1, 340 - 344

NEMETH, C. 1970
Bargaining and reciprocity.
Psychological Bulletin, 74, 297 - 308

NEWCOMB, T.M. & FELDMAN, V.A. 1969
The impact of college on students.
San Francisco

NISBETT, R.E. 1968
Cooperation
In: SILLS, D.L. (Ed.)
International Encyclopedia of the Social Science, Vol. 3, 384 - 390

NUNNALLY, J. 1960
The place of statisitics in psychology.
Educational and Psychological Measurement, 20, 641 - 650

O'DONELL Jr.,P.G. 1978
Evoking trustworthy behavior of children and adults in a Prisoner's Dilemma Game
Dissertation: University of Washington

OSGOOD, C. 1962
An alternative to war or surrender.
Urbana

OSKAMP, S. & PERLMAN, D. 1965
Factors affecting cooperation in a Prisoner's Dilemma Game.
Journal of Conflict Resolution, 9, 359 - 374

PAULSON, F.L. 1974
Teaching cooperation on television. An evaluation of Sesame Street social goal programs.
AVCR, 22, 229 - 246

PAWLOW, I.P. 1927
Conditioned reflexes.
London

PFLAUM, S.W. et al. 1980
Methods of teaching reading.
Evaluation in Education, 4, 121 - 122

PHILLIPS, B.N. & DeVAULT, M.V. 1957
Evaluation of research on cooperation and competition.
Psychological Reports, 3, 289 - 292

PILLEMER, D.B. & LIGHT, R.J. 1980
Synthesizing outcomes: How to use research evidence
from many studies.
Harvard Educational Review, 50, 176 - 195

POPPER, K.R. 1934
Logik der Forschung.
Tübingen

POPPER, K.R. 1963
Conjectures and refutations.
London

PRESBY, S. 1978
Overly broad categories obscure important differences
between therapies.
American Psychologist, 33, 514 - 517

RAPOPORT, A. 1967
A note on the index of cooperation for PDG.
Journal of Conflict Resolution, 11, 100 - 103

RAPOPORT, A. & CHENT, C. 1962
Experimental Games: A review.
Behavioral Science, 7, 1 - 37

RAPOPORT, A. & CHAMMAH, M.A. 1965
Prisoner's Dilemma. A study in conflict and cooperation.
Ann Abor

ROSENBERG, S. & HALL, R.L. 1958
The effects of differntial social feedback conditions
upon performance in dyadic teams.
Journal of Abnormal and Social Psychology, 57, 271 - 277

ROSENBERG, S. 1960
Cooperative behavior in dyads as a function of
reinforcement parameters.
Journal of Abnormal and Social Psychology, 60, 318 - 333

ROSENBERG, S. 1963
Influence and reward in structered two person interactions.
Journal of Abnormal and Social Psychology, 67, 379 - 387

ROSENHAN, D.L. 1972
Learning theory and prosocial behavior.
Journal of Social Issues, 28, 151 - 163

ROSENHAN, D.L. & WHITE, G.N. 1967
Observation and rehearsal as determinants of prosocial behavior.
Journal of Personality and Social Psychology, 5, 424 - 431

ROSENTHAL, R. 1969
Interpersonal expectations: effects of the experimenter's hypothesis.
In: ROSENTHAL, R. & ROSNOW, R.L. (Eds.): Artifact in behavioral research.
New York

ROSENTHAL, R. 1978
Combining results of independent studies.
Psychological Bulletin, 85, 185 - 193

ROSENTHAL, R. 1980
Combining probabilities and the file drawer problem.
Evaluation in Education, 4, 18 - 21

ROSENTHAL, R. & RUBIN, D.B. 1978
Interpersonal expectancy effects: the first 345 studies.
The Behavioral and Brain Sciences, 13, 377 - 415

SCHILLER, D., WALBERG, H.J., & HAERTEL, G.D. 1980
Quality of instruction
Evaluation in Education, 4, 119 - 120

SCHMIDT, F.L. & HUNTER, J.E. 1977
Development of general solution to the problem of validity generalization.
Journal of Applied Psychology, 62, 529 - 540

SCHMIDT, F.L. etal. 1979
Further tests of the SCHMIDT-HUNTER bayesian validity generalization procedure.
Personnel Psychology, 32, 257 - 281

SCHMITT, D.R. & MARWELL, G. 1971a
Taking and the disruption of cooperation.
Journal of the Experimental Analysis of Behavior, 15, 405 - 412

SCHMITT, D.R. & MARWELL, G. 1971b
Avoidance of risk as a determinant of cooperation.
Journal of the Experimental Analysis of behavior, 16, 367 - 374

SCHWAB, D.P., OLIAN-GOTTLIEB, J.D., & HENEMAN III, H.G. 1979
Between subject's expectancy theory research: A statistical
review of studies prediciting effort and performance.
Psychological Bulletin, 86, 139 - 147

SCODEL, A. et al. 1959
Some descriptive aspacts of two person non-zero-sum games.
Journal of Conflict Resolution, 3, 114 - 119

SEARS, R.R. 1951
A theoretical framework for personality and social behavior.
American Psychologist, 6, 476 - 483

SERBIN, L.A., TONICK, I.J., & STERNGLANZ, S.H. 1977
Shaping cooperative cross sex-play.
Child Development, 48, 924 - 929

SHAPIRO, A. & MADSEN, M.C. 1969
Cooperative and competitive bahavior of kibbutz and
urban children in Israel.
Child Development, 40, 609 - 617

SHAW, M.E. 1958
Some motivational factors in cooperation and competition.
Journal of Personality, 26, 155 - 169

SIDMAN, M. 1952
A note on functional relations obtained from group data.
Psychological Bulletin, 49, 263 - 269

SIDOWSKI, J.B., WYCKOFF, L.B., & TABORY, L. 1956
The influence of reinforcement and punishment in a minimal
social situation.
Journal of Abnormal and Social Psychology, 52, 115 - 119

SKINNER, B.F. 1963
Science and human behavior.
New York
deutsch: Wissenschaft und menschliches Verhalten (1973)
München

SKINNER, B.F. 1969
Contingencies of reinforcement - A theoretical analysis.
New York
deutsch: Die Funktion der Verstärkung in der Verhaltenswissenschaft (1974)
München

SKINNER, B.F. 1962
Two 'synthetic social relations'.
Journal of Experimental Analysis of Behavior, 5, 531 - 532

SLABY, R.G. & CROWLY, C.G. 1977
Modification of cooperation and aggression through teacher attention to children's speech.
Journal of Experimental Child Psychology, 23, 442 - 458

SLAVIN, R. 1980
Cooperative learning.
Review of Educational Research, 50, 315 - 342

SLAVIN, R.E. 1981
Synthesis of research on cooperative learning.
Educational Leadership, 38, 655 - 660

SLAVIN, R.E. 1983
Meta-Nonsens: Misuse of meta-analysis in educational research.
Paper presented at the Annual Convention of the American Educational Research Association, Montreal

SMITH, M.L. 1980
Sex bias in counseling and psychotherapy.
Psychological Bulletin, 87, 392 - 407

SMITH, M.L. & GLASS, G.V. 1977
Meta-analysis of psychotherapy outcome studies.
American Psychologist, 32, 752 - 760

SMITH, M.L. & GLASS, G.V. 1980
Meta-analysis of research on class size and its relationship to attitudes and instruction.
American Educational Research Journal, 17, 419 - 433

SMITH Jr., N.C. 1970
Replication studies: A neglected aspect of psychological research.
American Psychologist, 25, 970 - 975

SOLOMON, L. 1960
The influence of some types of power relationships and game strategies upon the development of interpersonal trust.
Journal of Abnormal and Social Psychology, 61, 223 - 230

SPINNER, H.F. 1977
Begründung, Kritik und Rationalität. Band 1.
Braunschweig

STAUB, E. 1973
Children's sharing behavior: Success and failure, the "norm of deserving" and reciprocity in sharing.
Paper presented at the symposion entitled "Helping and Sharing": Concepts of altriums and cooperation" at the meeting of the Society of research in Child Development, Philadelphia

STAUB, E 1978
Positive social behavior and morality. Social and personal influences (Vol. I).
New York

STAUB, E. 1979
Positive social behavior and morality. Socialization and development (Vol. II).
New York

STEELE, M.V. & TEDESCHI, J.T. 1967
Matrix indices and strategy choices in mixed-motive games.
Journal of Conflict Resolution, 11, 198 - 205

STEGMÜLLER, W. 1974
Probleme und Resultate der Wissenschaftstheorie und Anlytischen Philosophie.
Berlin, Heidelberg

STEPHENS, J.M. 1967
Making dependable use of published research: A proposed check list.
Journal of Educational Research, 61, 99 - 104

STERLING, T.D. 1959
Publication decisions and their possible effect: Interferences drawn from tests of siginificance - or vice versa.
Journal of American Statistical Association, 54, 30 - 34

STEVENSON, H.W. 1965
Social reinforcement of children's behavior.
In: LIPSETT, L.P. & SPIKER, C.C. (Eds.), Advances in child development.
New York

STRAUSS, S. 1969
Guidelines for analysis of research reports.
Journal of Educational Research, 63, 165 - 169

SWINGLE, P.G. & COADY, H. 1967
Effects of the partner's abrupt strategy change upon the subject's rsponding in the Prisoner's Dilemma.
Journal of Personality and Social Psychology, 8, 160 - 165

TEDESHI, J.T., HIESTER, D., & GAHAGAN, J.P. 1969
Matrix values and the behavior of children in the Prisoner's Dilemma Game.
Child Development, 40, 517 - 527

THEROUX, S.S. 1974
The effect of modeling of cooperation in young children.
Dissertation: University of Massachusetts

THIBAUT, J.W. & KELLEY, H.H. 1959
The socal psychology of groups.
New York

THOMPSON, P.R. 1980
"And who is my neighbour?" An answer from evolutionary genetics.
Social Science Information, 19, 341 - 384

THORNDIKE, E.L. 1898
Animal intelligence: An experimental study of associative processes in learning.
Psychological Review Monographs Supplement, 5, 551 - 553

THORNDIKE, E.L. 1911
Animal intelligence.
New York

TORRANCE, E.P. 1971
"Structure" can improve the group behavior of five-year-old children.
Elementary School Journal, 72, 102 - 106

TUKEY, J.W. 1977
Exploratory data analysis.
Reading

UGUROGLU, M.E. & WALBERG, H.J. 1980
Motivation
Evaluation in Education, 4, 105 - 106

VINACKE, W.E. 1969
Variables in experimental games.
Psychological Bulletin, 71, 293 - 313

VOGLER, R.E. 1968a
Possibility of artifact in studies of cooperation.
Psychological reports, 23, 9 - 10

VOGLER, R.E. 1968b
Awareness and the oparant conditioning of a cooperative response.
Journal of Psychology, 69, 117 - 127

VOGLER, R.E., MASTERS, W.M., & MORRILL, G.S. 1970
Shaping cooperative behavior in young children.
Journal of Personality, 74, 181 - 186

VOGLER, R.E., MASTERS, W.M., & MORRILL, G.S. 1971
Extinction of cooperative behavior as a function of acquisition by shaping or instruction.
Journal of Genetic Psychology, 119, 233 - 240

WALBER, H.J. & HAERTEL, E.H. 1980
Research integration: An introduction and overview.
Evaluation in Education, 4, 5 - 10

WALLER, M. 1973
Kooperatives Verhalten im Verständnis von Kindern als symmetrischer Austauschprozeß: Stärke und Bedingungen der Ausbildung der Norm des symmetrischen Austausches in den Verhaltenserwartungen 4 - 9jähriger Kinder.
Zeitschrift für Sozialpsychologie, 4, 1973, 51 - 61

WALSTER, E., BERSCHEID, E., & WALSTER, G.W. 1970
Reaction of an exploiter to the exploited: Compensation, justification, or self-punishment?
In: BERKOWITZ, L. & MACAULAY, J.R. (Eds.), Altruism and helping behavior
New York

WALSTER, E. & PILIAVIN, J.A. 1972
Equity and the innocent bystander.
Journal of Social Issues, 28, 165 - 189

WALSTER, E., BERSCHEID, E., & WALSTER, G.W. 1973
New directions in equity research.
Journal of Personality and Social Psychology, 25, 151 - 176

WALSTER, E. & WALSTER, G.W. 1975
Equity and social justice.
Journal of Social Issues, 31, 21 - 43

WALSTER, E., WALSTER, G.W., & BERSCHEID, E. 1978
Eqity: Theory and research.
Boston

WAY, J.G. & LARIMORE, D.L. 1973
A joinder on signifying significant significance.
American Psychologist, 28, 361 - 362

WAXMAN, H.C. & WALBERG, H.J. 1980
Teacher effectiveness.
Evaluation in Education, 4, 123 - 124

WEINGARTNER, K. & MECHNER, F. 1966
The contingency as an independent variable of social interaction.
In: VERHAVE (ed.), The experimental analysis of behavior.
New York

WESTMEYER, H. 1973
Kritik der psychologischen Unvernunft.
Stuttgart, Köln

WESTMEYER, H. 1979
Wissenschafttheoretische Grundlagen der Einzelfallanalyse.
In: PETERMANN, F: & HEHL, F.-J. (Hrsg.), Einzelfallanalyse.
München

WHITE, K. 1976
The relationship between socioeconomic status and academic achievement.
Dissertation: University of Colorado at Boulder

WILKE-BIRKENHAUER, U. 1976
Experimentelle Untersuchungen zur Analyse kooperativen Verhaltens bei Kindern.
Dissertation: Universität Düsseldorf

WILLIAMS, P.A., HAERTEL, G.D., & WALBERG, H.J. 1980
Media.
Evaluation in Education, 4, 109 - 110

WILLIAMS, P.A: et al. 1982
The impact of leisure time television on school learning:
A research synthesis.
American Educational Research Journal, 19, 19 - 50

WILSON, W. 1971
Reciprocation and other techniques for inducing cooperation
in the Prisoner's Dilemma Game.
Journal of Conflict Resolution, 15, 167 - 165

WINCH, R.F. & CAMPBELL, D.T. 1969
Proof? No! Evidence? Yes! The significance of tests of
siginificance.
American Sociologist, 4, 140 - 143

WINER, B.J. 1971
Statistical principles in experimental design.
New York

WISPE, L.G. 1972
Positive forms of social behavior: An overview.
Journal of Social Issues, 28, 1- 19

WOLF, E. 1972
über das Zusammenwirken der kognitiven und emotionalen
Komponenten von Einstellungen zu Kooperationspartnern.
Zeitschrift für Psychologie, 180/181

WOLF, W. 1973
Ein Flußdiagramm als Hilfe zur Beurteilung empirischer
Untersuchungen.
Zeitschrift für Pädagogik, 19, 63 - 76

YARROW, M.R., SCOTT, P.M., & WAXLER, C.Z. 1973
Learning concern for others.
Develop Mental Psychology, 8, 240 - 261

ZETTERBERG, H.L. 1954
On theory and verification in sociology.
New York

Anhang

ANHANG 1

Kodiersystem

..... Studiennummer 01
..... laufende Nummer in der Meta-Analyse 02
..... laufende ES-Nummer 03
..... laufende ES-Nummer pro Studie 04
..... Kodier-Datum 05
..... Wiederholung der Kodierung 06
..... Autor(en) 07
.....
.....
.....
..... Anzahl der Autoren 08
..... Publikationsjahr 09
..... Publikationstyp 10
Journal = 1 Dissertation = 2 Buch = 3

..... Publikationsumfang (Seiten) 11
..... Anzahl der Experimente 12
..... Einzelfall = 1 Gruppen = 2 13
..... Anzahl berechneter Effektgrößen 14

..... Explikation der Fragestellung 15
ja = 1 nein = 0

..... Kategorie der Fragestellung 16
(Explikation a post.)

..... Explikation der Problemstellung 17
ja = 1 nein = 0

..... Kategorie der Problemstellung 18
(Kalssifikation a post.)

..... Bedeutung der Studie 19
gering = 1 hoch = 5

..... Referenztheorie für die Unabhängigen Variablen 20
Sozialkognitive Entwicklungstheorie = 1
Altersbezogene Entwicklungstheorie = 2
Operante Konditionierung = 3
Sozialkoginitive Lerntheorie = 4
Austauschtheorie = 5
Feldtheorie = 6

..... Übereinstimmung zwischen 17 und 20
ja = 1 nein = 0 21

.....	Referenzkonzept (Klassifikation a post.)	22
.....	Literaturbenennung zu Referenzen gering = 1 hoch = 5	23
.....	Anzahl inhaltlicher Hypothesen	24
.....	Anzahl statistischer Hypothesen	25
.....	Explikation des Konstrukts ja = 1 nein = 0	26
.....	Kategorie des Konstruks (Klassifikation a post.)	27
.....	Operationale Definition der UV ja = 1 nein = 0	28
.....	Extraktion der Treatentbedingungen (Klassifikation a post.)	29
.....	Reward	30
.....	Temptation	31
.....	Sucker's pay off	32
.....	Punishment	33
.....	Dauer der Treatmentapplikation Minuten	34
.....	Dauer der Treatmentapplikation Züge	35
.....	Definition des AV-Konstrukts ja = 1 nein = 0	36
.....	Treatment x AV Treatment und Meßvorgang sind konfundiert = 0 Meßvorgang ohne vorheriges Treatment = 1 Treatment plus Meßvorgang = 2	37
.....	Klassifikation der operationalen AV Interd-resp.shar.-nonsoc-forced = 1 Interd-resp.shar.-nonsoc-alt.resp = 2 Interd-resp.shar.-soc-forced = 3 Interd-resp.shar.-soc-alt.resp = 4 Interd-resp.exch.-nonsoc.-forced = 5 Interd-resp.exch.-nonsoc.-alt.resp = 6	38

Interd-resp.exch.-soc.forced = 7
Interd-resp.exch.-soc.-alt.resp. = 8
Dependent-resp.exch.-nonsoc-forced = 9
Dependent-resp.exh.-nonsoc.-alt.resp. = 10
Dependent-resp.exch.-soc.-forced = 11
Dependent-resp.exch.-soc.-alt.resp = 12
Nichtnullsummenspiele = 14
Beobachtung = 20

..... AV-Aggregat 39
Episode = 1 Individualreaktion = 2

..... Reliabilität 40

..... Definition der Population 41
ja = 1 nein = 0

..... Stichprobengröße 42

..... Anzahl genannter Stichprobenmerkmale 43

..... Grad der Kompetition 44
niedrig = 1 mittel = 2 hoch = 3

..... ökographie 45
Stadt = 1 Land = 2

..... Alter der Probanden 46

..... Geschlecht der Probanden 47
männl. = 1 weibl. = 2 konfundiert = 3

..... Sozioökonomischer Status 48
Unter = 1 Mittel = 2 obere Mittel = 3 Ober = 4

..... Ziehung der Stichprobe 49
Klumpen = 1 Random = 2 Freiwillig = 3
Zugewiesen = 4 Sonstiges = 5

..... Gruppenzuweisung 50
Matching = 1 Random = 2

..... Interne Validität 51

..... Externe Validität 52
niedrig = 1 mittel = 2 hoch = 3

..... Konstruktvalidität 53

..... Validität des statistischen Schlusses 54

..... Datenbasis 55

Prozentwerte = 1		Non-parametrische Stat.	= 2
Mittelwerte = 3		parametrische Statistiken	= 4
SD aus Einweg= 5		SD aus Mehrweg	= 5
Korrelation = 7		empirischer p-Wert	= 8
kritischer p = 9		Sonst	=10

..... Quelle des Mittelwertes 56

unjust. Postwert	= 1	covarianz-just. Wert	= 2
Residualwert	= 3	Vortest-Nachtest-Diff	= 4
Rest	= 5		

..... AM(T) Vor 57
..... SD(T) Vor 58
..... AM(C) Vor 59
..... SD(C) Vor 60
..... AM(T) Nach 61
..... SD(T) Nach 62
..... AM(C) Nach 63
..... SD(C) Nach 64

..... Effektgröße 65

..... Prozentwert C 66

..... n Treatmentgruppe 67

..... n Kontrollgruppe 68

..... Analyseeinheit 69
Individuum = 1 Dyade = 2 Triade = 3 gt 3 = 4

..... Übereinstimmung Diskussion - Ergebnisse 70
ja = 1 nein = 0

..... Letzte Effektgröße der Studie 71
ja = 1 nein = 0

..... AM(ES) der Studie 72

Anhang II

Inhalte der Kategorien zur Fragestellung

Studiencode	Inhalt der Fragestellung
01	KONTINGENTE BEKRÄFTIGUNG
0101	Führt gruppenkontingente Bekräftigung kooperativer Episoden zur Erhöhung kooperativen Verhaltens?
0202 1101	Bewirkt soziale Verstärkung eine Zunahme korrelierter sozialer Verhaltensweisen?
0203	Wirken Verstärkungskontingenzen kooperationserhöhend oder allgemein reaktionserhöhend?
0701	Wirken verschiedene Revenuen auf kooperatives Verhalten unterschiedlich?
0801	Beeinflußt soziale Verstärkung das Wahlverhalen im PDG?
0803 1401 1801 2201	Welche Verstärkungsbedingungen beeinflussen Kooperation?
1301 1304	Führt guppenkontigente gegenüber individualkontingenter Verstärkung zu unterschiedlichen Kooperationsraten?
1302	Führt kontingente Belohnung kooperativer Reaktionen bei paralleler Bestrafung kompetitiver Reaktionen zu höheren Kooperationseffekten?
1310	Wirken positive Vorerfahrungen und ein hohes Verstarkungsniveau kooperationsfördernd?
1901 1902	Wirkt kontingente Aufmerksamkeit kooperationsfördernd?
2202	Führt die Kenntnis der Kontingenz zu höherer Kooperation?
2203	Wirken Bekräftigungskontingenzen für Kooperation auch bei 4-6jährigen Kindern ?

02 MODELLIERUNG KOOPERATIVEN VERHALTENS

0201 Erhöht die modellierte Kontingenz das Reaktionsverhalten im PDG?

0205 Bewirkt die stellvertretende Darstellung einer TFT-Strategie höhere Kooperationsraten?

0206
0301
1601 Wirken prosoziale Filme gegenüber aggressiven Filmen kooperatinsfördernd?

0601 Wirkt die stellvertretende, externale gegenüber selbstinduzierter Bekräftigung kooperationsfördernd?

0804 Welche der beiden Theorien - PIAGET oder BANDURA - erklären Kooperation 'besser'?

0901 Wirkt stellvertretendes Lernen bei kooperativen Sequenzen?
2002
2003

1203 Wirkt die GRIT-Strategie kooperationsfördernd?

03 INSTRUKTION ZUR STRATEGIE

1903 Welche motivationalen Faktoren beeinflussen Kooperation?

1904 Bewirken reaktionssequentielle Strategien eine Veränderung in der Höhe der Kooperationsrate?

2001 Erleichtert die Transparenz der Aufgaben- oder Gruppenstruktur Kooperation?

2204 Unterscheiden sich Shaping und Instruktion in der Akquisition von Kooperation?

04 VARIATION DER MATERIXINDIZES

0702 Führt die Variation der Matrixindices zur Variation der
1307 Auftrittswahrscheinlichkeit kooperativen (Wahl-) Ver-
1501 haltens?
2301

05 HINWESIE ZUM VERHALTEN DES KO-AKTEURS

0703 Bewirkt Information über die Ähnlichkeit zwischen Probanden und Ko-Akteur eine höhere Kooperationsrate?

1001
1201 Prädisponiert die Information über die Kooperativität des Mitspielers?

1305 Beeinflußt die Koopertionsrate des Ko-Akteurs das kooperative Verhalten der Probanden?

1501 Wirken Gruppennormen auf die Reaktionsrate kooperativen Verhaltens

06 AUFGABENSTRUKTUR

0204 Welchen Einfluß übt das Aufgabenniveau aif kooperatives Verhalten aus?

1204 Wie wirkt die Aufgabenstruktur auf Kooperation?

2001 Erleichtert die Transparenz der Aufgabenstruktur kooperatives Verhalten ?

07 RESIDUALKATEGORIE

1202 Kann Kooperation wie Individualverhalten gemessen werden?

1309 Welche Grundfertigkeiten und welche Anreize fuhren zu Kooperation?

1502 Welchen Einfluß übt Vertrauen und sich selbst erfullende Prophezeihung auf Kooperation aus?

Anhang III

Inhalte zum Referenzkonzept

Studien code	Kategorie	Kategorieninhalte
0101	01	kontingente Gruppenverstärkung
0201	05	Stellvertretende Verstärkung (Modellierung)
0202	03	soziale Verstärkung (intermittierend)
0203	01	Bekräftigungskontingenzen
0204	07	Vertrauen in Form 83% C-Wahlen
0205	05	Stellvertretende Verstärkung (Modellierung)
0206	06	Modellierung mittels Fernsehen
0301	06	Modellierung der Reaktion
0601	06	Selbstinduzierte Bekräftigung (Modellierung)
0701	01	Revenuen-Niveaus
0702	08	Matrixindices
0703	08	Ziel-Orientierung
0801	01	Soziale Bekräftigung
0802	09	Nachrichtenmerkmale
0803	01	Kontingente Bekräftigung
0804	09	Alter der Probanden
0901	06	Modellierung interferenten Verhaltens
1001	09	Attribution
1101	01	Soziale Verstärkung
1201	07	Verläßlichkeit kooperativer (C-) Wahlen
1202	02	differentielle Verstärkung
1203	07	Vertrauen durch GRIT
1204	08	Zielstruktur
1301	01	Vorerfahrung an Verstärkung
1302	01	Individuelle Bekräftigung für kooperative Episoden
1303	01	Gruppen- gegenüber individueller Kontingenz
1304	01	Gruppenkontingente Bekräftigung
1305	08	Gewinnorientierung x Matrixindices
1306	01	Bekräftigung durch Auszahlungsanzeige
1308	05	Stellvertetende gegenuber direkter Bekraftigung
1401	02	Diskrimination von Verstärkungskontingenzen
1501	08	Gruppennormen x Matrixindices
1502	07	Vertrauen
1601	05	Stellvertretende Kontingenzen (Modellierung)
1801	01	Soziale Bekräftigung
1901	01	Soziale Bekräftigung (Aufmerksamkeit)
1902	01	Soziale Bekräftigung (Aufmerksamkeit)
1903	08	Motivinduktion durch Instruktion
1904	06	Modellierung von Vertrauen
2001	09	Aufgaben- und Organisationsstruktur
2002	05	Stellvertretende gegenuber direkter Bekräftigung
2003	06	Reaktionsmodellierung
2201	01	Bekräftigungskontingenz

2204	04	Transparenz der Kontingenz (Awareness)
2203	01	Latenz der Konsequenz
2204	02	Shaping und Instruktion
2301	07	Vertrauen durch Strategie des Ko-Akteurs

Die verschiedenen Konstrukte, auf die sich die Studien in ihren Hypothesen bezogen, wurden wie folgt kategorisiert:

01	Kontingente Bekräftigung für gebende Reaktionen
02	Differentielle Verstärkung
03	Kontingente Verstärkung für prosoziales Verhalten
04	Wahrnehmung der Kontingenz
05	Modellierung der Kontingenz
06	Modellierung der Reaktion
07	Vertrauen
08	Motivinduktion
09	Restkategorie

Anhang IV

Klassifikation der Problemstellung

Unter der Kategorie 'Problemstellung' wurde solche Information zusammengefaßt, die den Zusammenhang zwischen zu testender Variable und dem Theoriekern herstellt. Die aus den Studien extrahierte Information wurde a posteriori aggregiert.

Studiencode	Kategorieninhalte
01	VERSTÄRKUNGSPARADIGMA
0101	Bildet die gruppenkontingente Verstärkung kooperativer Episoden eine hinreichende Bedingung?
0201	Im PDG sind die immantenten Kontingenzen unbekannt
0601	Welche Kontingenzen und welche Form der Modellierung wirkt kooperationsfördernd?
0803	Führen Verstärkungsstrategien zu kooperativem Verhalten
1304	Welche Komplexität muß eine Kontingenz aufweisen, um kooperatives Verhalten in Triaden zu evozieren?
1308	Läßt sich durch Verstärkung bei amerikanischen Kindern Kompetition besser implimentieren als Kooperation?
1401	Erklärt eine Verstärkungsanalyse die widersprüchlichen Kooperationsergebnisse?
1801	Welche Verstärkungs-Parameter führen zu Kooperation?
1901	Erhöht kontingente Aufmerksamkeit sowohl das Zielverhalten wie auch korrelierte Verhaltensweisen
2202	Wirkt die Kenntnis der Kontingenz auf komplexe Reaktionssequenzen als hinreichnde Bedingung
02	INTERFERNZFAKTOREN
0204 1501	Welche Antezendenzen führen zu hohen Kooperationsraten über eine große Anzahl von Zügen
0703	Wirkt Information zur Situation interferenzreduzierend?
0801 2301	Welche Variablen reduzieren im PDG die Interferez?

1302 Welche Risiko-Faktoren behindern Kooperation?

1305 Welchen Einfluß üben konditionale Strategien aus?

1902 Wirkt rollenspezifische Sozialisation in gemischt-geschlechtlichen Dyaden kooperationshindernd?

1903 Liefert die 'Effektivität der Zielerreichung' einen Erklärungsansatz

2001 Verhindert bei jungen Kindern die Desorientierung sozialer Organisationen (in Gruppen) eine konstruktive Orientierung?

03 MODELLIERUNG

0205 Wirkt hohe Ich-Beteiligung störend auf die Akquisition?

0206 Besteht ein Zusammenhang zwischen Kompetition und Aggression gegenüber Kooperation und Prosozialität und ihrer Modellierung durch Filme?

0301
2003 Kann mittels Modellierung Kooperation gefördert werden?

0601 Welche Modellierungsstrategien wirken kooperationsfördernd?

0804 Welche Theorie (PIAGET gegenüber BANDURA) erklärt Kooperation 'besser'?

1601 Kann Sesame Street Sozialverhalten beeinflussen?

04 VERSTÄRKUNG KORRELIERTER VERHALTENSWEISEN

0202
1901 Bewirkt soziale Bekräftigung eine Veränderung ähnlicher Verhaltensweisen

0203
1202 Können frei-operante Konditionierungen auf soziale Interaktionen angewandt werden?

0901 Kann durch Kooperation Aggression kontrolliert werden?

2201 Werden durch auf Kooperation bezogene Verstärkungsbedingungen auch nicht-kooperative Verhaltensweisen verstärkt?

05 VARIATION DER MATRIXINDICES

0701 Impliziert das PDG immanent wenig Anreize zur Kooperation?

0702 1307 Beeinflußt die Variation der Matrixwerte die Kooperationsrate?

1301 Beeinflußt die Erhöhung der Ergebnisdifferenz die Reaktionsrate?

1306 Hilft eine MDG-Matrix, das Motiv (des Ko-Akteurs) leichter zu erkennen?

06 VERHALTENSMERKMALE DES KO-AKTEURS

1001 Beeinflußt die Attribution des Ko-Akteu Kooperation?

1201 Ist Rollenübernahme als Antizipation der Verhaltenstendenz des Ko-Akteurs bei Kindern zwischen 9 und 13 Jahren möglich?

1203 Bewirken verschiedene Strategien des Ko-Akteurs unterschiedliche Kooperationstendenzen

1502 Hilft induziertes Vertrauen beim Aufbau koopertiver Reaktionen

1904 Welchen Einfluß üben reaktionssequentielle Strategien aus?

07 ENTWICKLUNGSDETERMINANTEN

0804 Welche Theorie erklärt die Genese sozialen Verhaltens (PIAGET vs BANDURA) ?

1204 Wie können Motive und soziale Fertigkeiten bei Kindern entwickelt werden?

1309 Welche Faktoren beeinflussen die Entwicklung von Kooperation bei Kindern?

2001 Ist kindliche Aggression durch Desorientierung in sozialen Strukturen bedingt?

2203 Wirkt kontingente Belohnung auch bei jungen Kindern?
1101
2203

1901 Welchen Einfluß üben bei Kindern demographische Daten aus?

08 DISKRIMINATIONSLERNEN

2204 Welche Modifikationsstrategie (Instruktion versus Shaping) wirkt bei der Ausformung kooperativen verhaltens?

09 RESTKATEGORIE

0802 Welche Wirkung üben Nachrichtensendungen auf Formen des sozialen Verhaltens aus?

2002 Kann die gesellschaftliche Erwartung an das sozial-gestörte Kind durch Erhöhung seiner kooperativen Kompetenz verändert werden?

Anhang V

Funktionale Analyse der Effektgrößen-bezogenen Treatments

Studien code	Nummer in MA		Funktionale Analyse
0201	01	01	Modell (A) = 80% R(G) + Ko-Akteur (I) = 80% R(G) ⟶ 60% K(+) für (A)R(G)
		02	Modell (A) = 80% R(G) + Ko-Akteur (II) = 80% R(N) ⟶ 80% K(-) für (A)R(G)
		03	MOdell (A) = 80% R(N) + Ko-Akteur (III) = 80% R(G) ⟶ 80% K(+) für (A)R(N)
		04	Modell (A) = 80% R(N) + Ko-Akteur (IV) = 80% R(N) ⟶ 60% K(-) für (A)R(N)
		05	Proband mit I / TFT
		06	Proband mit II/TFT
		07	Proband mit III/TFT
		08	Proband mit IV/TFT
		09	Proband mit neutrlem Ko-Akteur
0101	02	01	(A)R(G) + (B)R(G) ⟶ 1 (A,B) K(+)
0202	03	01	K(+)/soz ⟶ PSV
		02	K(+)/soz + priming ⟶ PSV
1310	04	01	Vorspiel: Ko-Akteur zeigt R(G)
		02	Vorspiel: Ko-Akteur zeigt R(N)
		03	Meßvorg.: Revenuen hoch
		04	Meßvorg.: Revenuen niedrig
0701	05	01	Revenuen: imaginär
		02	Revenuen: 1-2 cent
		03	Revenuen: 5-10 cent
		04	Revenuen: 10-20 cent
		05	Revenuen: Dollar
		06	Anfangs-R(G) niedrig-----mittel----hoch
0702	06	01	Matrixindizes
1001	07	01	B(D): Du bist kooperativ
1101	08	01	K(+) für Sozialkontakt
		02	K(+) = Cents für Sozialkontakt
		03	K(+) = Süßigkeiten für Sozialkontakt
		04	K(+) = Zuwendung der anderen Kinder

1201	09	01	p(K(+)/R(G) = .1
		02	p(K(+)/R(G) = .5
		03	p(K(+)/R(G) = .9
		04	Revenuen = 5 Dollar Spielgeld
		05	Revenuen = 10 Dollar Spielgeld
		06	Revenuen = 20 Dollar Spielgeld
		07	Strategie = 50% Kooperation (unkontingent)
		08	Strategie = 90% kooperation (unkontingent)
		09	männlich
		10	weiblich
1301	10	01	(A)R(G) + (B)R(G) ⟶ (A)K(+) + (B)K(+) Gruppenkontingenz
		02	(A)R(G) + (B)R(G) ⟶ (B)K(+) Episode Individualkontingent
		03	männlich
		04	weiblich
		05	Stadt
		06	Land
1302	11	01	(A)R(G) + (B)R(G) ⟶ (A)K(+) + (B)K(+)/ R(I) wurde nicht belohnt (335 Züge)
		02	(A)R(G) oder (A)R(I) + (B)R(G) oder (B)R(I) ⟶ (A)K(+) + (B)K(+)/ R(G) wurde nicht belohnt (+ 135 Züge)
		03	Revenuen (01, 02) = .6 cent/ K(+)
		04	(A)R(G) + (B)R(G) ⟶ (A)K(+) + (B)K(+)/ (A)R(I) + (B)R(I) ⟶ (A)K(+) + (B)K(+)/ K(+) für R(I) lt K(+) für R(G): .1 lt .6 cent
		05	wie 04 jedoch K(+) für R(I) gt K(+) für R(I): .3 gt .2 cent
		06	Risiko: A oder B konnten vom Ko-Akteur pro Episode .1 cent nehmen: (A)R(G) + (B)R(G) ⟶ (A)K(+ = .3c) + (B)K(+ = .3c)/ R(N) = K(- = .1c) für B/A + (A)R(I) + (B)R(I) ⟶ (A)K(+ = .2c) + (B)K(+ = .2c)
		07	Risiko hoch: statt Cent Dollar
		09	lange Spieldauer für 07
		10	Revenuen für R(G) auf .9 cent
		11	Risiko ohne Nutzen für den Ko-Akteur: R(N) wurde nicht belohnt
		12	R(N) konnte für jede Reaktion gtedrückt werden
		13	Kommunikation
		14	verzögerte Kommunikation

1202	01	(A)R(G) + (B)R(G) ⟶ (A)K(+) + (B)K(+) +
		(A)R(I)/(B)R(I) ⟶ (A)K(-) + (B)K(-) +
		S(D): Verhalten des Ko-Akteurs +
		Latenz lt .5'' +
		Alter = 18 Jahre
	02	wie 01/ Alter = 12 J/hre
	03	wie 01/ Interaktionspartner waren Freunde
	04	wie 01/ Interaktionspartner sind kompetitive Brüder
	05	(A)R(G) + (B)R(G) ⟶ (B)K(+) Episode
	06	wie 05/ Freunde
	07	wie 05/ Geschwister
	08	wie 01/ mechanische Anzeige
	09	wie 01/ Gesichtskontakt
1304	01	(A)R(G) + (B)R(G) + (C)R(G) ⟶ indiv. K(+) +
		(A)R(I) + (B)R(I) + (C)R(I) ⟶ indiv. K(+) /
		K(+) für R(I) lt K(+) für R(G) /
		Revenuen: 100 Punkte = 1 cent
	02	wie 01: Revenuen = 10000 Punkte = 1 cent
1501	01	S(D): Koop. = rationale Strategie
		Ko-Akteur spielt TFT
		S(D): Diskussion über kompetitive Norm
	02	wie 01,
		S(D): Diskussion über kooperative Norm
	03	wie 01
		keine Diskussion
	04	wie 01
		keine 100%ige TFT-Strategie
	05	wie 01
		andere Matrixwerte
	06	wie 01: 1. Drittel Spielverlauf
	07	wie 01: 2. Drittel Spielverlauf
	08	wie 01: 3. Drittel Spielverlauf
	09	wie 01: Anzeige der Gewinndifferenz
	10	wie 09: männlich
	11	wie 09: weiblich
	12	wie 09: gemischt-geschlechtlich
	13	wie 09: Änderung der Matrixwerte
	14	wie 09: neutrale Instruktion
	15	wie 09: 1. Drittel Spielverlauf
	16	wie 09: 2. Drittel Spielverlauf
	17	wie 09: 3. Drittel Spielverlauf

1305	16	01	Own-Gain/ Reaktionskontingenz des Ko-Akteurs
		02	Own-Gain/ Reaktions-Nonkontingenz des Ko-A.
		03	Relative-Gain/ Reaktionskontingenz
		04	Relative-Gain/ Reaktions-Nonkontingenz
1601	17	01	Modellierung: Reaktions-Sharing + S(D) = Koop.
1801	18	01	Shaping: p(K(+))Koop) = .5
1901	19	01	K(+) auf R(G)/ K(+) = sozial
		02	K(+) auf aggresive Reaktion/ K(+) = sozial
1902	20	01	K(+) auf R(G)/K(+) = sozial
1903	21	01	S(D): this is a cooperative situation: working together
		02	S(D): this is a competitive situation: your own individueal target5 is...
		03	S(D): neutrale Instruktion
2001	22	01	S(D): Koop
		02	wie 01 + Führung
2201	23	01	(A)R(G) + (B)R(G) ⟶ 1 (A,B)K(+)
0203	24	01	(A)R(G) + (B)R(G) ⟶ (A)K(+) + (B)K(+)
0703	25	01	S(D): falsche Information über Orientierung
		02	S(D): keine Information
		03	S(D): korrekte Information über Orientierung
2202	26	01	(A)R(G) + (B)R(G) ⟶ 1 (A,B)K(+) Bewußtheit der Kontingenz
		02	wie 01 Unkenntnis der Kontingenz
2203	27	01	(A)R(G) + (B)R(G) = 1 (A,B)K(+)
2204	28	01	(A)R(G) + (B)R(G) ⟶ 1 (A,B)K(+)
		02	wie 01 + Instruktion
0801	29	01	(A)R(G) ⟶ (A)K(+) + S(D): Ankündigung von Koop + % Verwirklichung von Koop Ankündigung zu Verwirklichung = 75 : 25
		02	wie 01 Ankündigung zu Verwirklichung = 25 : 75

0204	30	01	(B)R(G) = 83% (unkontingent)
		02	(B)R(G) = 17% (unkontingent)
0802	31	01	positive Nachrichten
		02	schlechte Nachrichten
0803	32	01	60 von 85 (B)R(G) (unkontingent)
		02	85 von 85 (B)R(G) (unkontingent)
1306	33	01	hohe Revenuen
		02	niedrige Revenuen
		03	eigene Anzeige
		04	eigene und fremde Anzeige
		05	Spielzuge
1307	34	01	Matrixindices
0205	35	01	Modellierung: Naive Person gegen TFT (= 45% R(G)
		02	Matrixindizes
1203	36	01	Vor = 50% R(G)/Zufall + GRIT
		02	Vor = 50% R(G)/Zufall + TFT
		03	Vor = 50% R(G)/Zufall + R(N)
		04	50% R(G)/Zufall
		05	01 + 100% R(G)
		06	02 + 100% R(G)
		07	03 + 100% R(G)
		08	04 + 100% R(G)
2301	37	01	(A)R(G) → (B)R(G) + (A)R(N) → 2 (B)R(N)
		02	TFT
		03	nach 3 (A)R(N)/TFT → (B)R(G)
		04	nach 2 (A)R(N)/TFT → (B)R(G)
		05	Vor: 100% R(N) + 01
		06	Vor: 100% R(N) + 02
		07	Vor: 100% R(N) + 03
		08	Vor: 100% R(N) + 04
		09	Vor: 87% R(N) + 01
		10	Vor: 87% R(N) + 02
		11	Vor: 87% R(N) + 03
		12	Vor: 87% R(N) + 04

0206	38	01	Modellierung: Mr. Roger's Nighborhood
		02	Modellierung: Speed Racer
		03	Modellierung: neitraler Film
		04	Meßinstrument
0601	39	01	Modellierung: (A)R(G) + (B)R(G) ⟶ 1 (A,B)K(+)/ K(+) = verbales Lob durch Erwachsenen
		02	wie 01 K(+) = verbales Lob durch den Ko-Akteur
1502	40	01	(E)S(D): Ich erwarte, du spielst kooperativ + (E)R(G) = 66%
		02	(E)S(D): Ich erwarte an, du wirst kompetitieren + (E)R(G) = 32%
		03	(E)R(G) = 66%
		04	Matrixindices
		05	(K)S(D): Ich nehme an, du wist kooperieren + (K)R(G) = 40%
		06	(K)S(D): Ich nehme an, du wirst kompetitieren + (K)R(G) = 14.5%
		07	(K)R(G) = 40%
		08	(E)R(G) lt 40%
		09	(K)S(D)/K: Ich nehme an, du spielst kooperativ + (K)R(G) = 29%
		10	(K)S(D)/K: Ich nehme an, du spielst kompetitiv + (K)R(G) = 30.7%
		11	(K)R(G) = 29%
		12	(K)R(G) = 30.7%
1401	41	01	S(D): play in a cooperative manner + Individualkontingenz
		02	S(D): play in a competitive way + K(+) nach Nullsumme
0901	42	01	Modellierung R(G) mit TFT + verbale Explikation
		02	Modellierung R(G) mit TFT - verbaler Explikation
1204	43	01	S(D): Instruktion = Kompetitiv
		02	S(D): Instruktion = Kooperativ
		03	02 + 01
		04	01 + 02

1308	44	01	S(D) = Koop + (A)R(G) + (B)R(G) ⟶ (A)K(+) + (B)K(+)
		02	S(D) = Komp + (A)R(N) + (B)R(G) ⟶ (A)K(+) + (B)K(-)
		03	01 als Film
		04	02 als Film
		05	01 bis 04 mit Matrix II : K(+) für Koop
		06	01 bis 04 mit Matrix III : K(+) für Komp
		07	01 bis 04 mit Matrix IV : K(+) für Koop + Komp
1309	45	01	S(D): Funktion des Spiels ⟵ Individualkontingenz für reziproke Periode + Interakteure sind bekannt
		02	wie 01 Interkteure kennen sich nicht
0804	46	01	Modellierung R(G) durch Individualkontingenz
		02	Modellierung R(N) durch Individualkontingenz
2002	47	01	Modellierung kooperativer Perioden
2003	48	01	Modellierung gemeinsamer Aktion
		02	Modellierung gemeinsamer Planung
		03	Modellierung 02 + 01
0301	49	01	Modellierung: Filmsequnzen mit hohem Koop aus Sesame Street
1904	50	01	100% R(G) (aktionssequentiell) + Matrix 1
		02	100% R(G) (aktionssequentiell) + Matrix 2
		03	100% R(G) (aktionssequentiell) + Matrix 3
		04	100% R(G) (reaktionssequentiell) + Matrix 1
		05	100% R(G) (reaktionssequentiell) + Matrix 2
		06	100% R(G) (reaktionssequentiell) + Matrix 3
		07	TFT + Matrix 1
		08	TFT + Matrix 2
		09	TFT + Matrix 3
		10	Matrix 1
		11	Matrix 2
		12	Matrix 3

LEBENSLAUF

Als Sohn der Eheleute Ulrich und Helga Tries, geborene Littman wurde ich am 4. 10. 1948 in Oberhausen (Rheinland) geboren. In Mülheim a. d. Ruhr wuchs ich auf und besuchte dort die Volksschule von 1955 bis 1963. Nach zweijährigem Besuch einer Privatschule in Duisburg absolvierte ich zwischen 1965 und 1968 eine Lehrausbildung am Max-Planck-Institut für Kohlenforschung in Mülheim a. d. Ruhr. Dort arbeitete ich ein weiteres Jahr. Zwischen 1969 und 1970 besuchte ich die Fachoberschule und erwarb die Fachhochschulreife. Von 1972 bis 1975 studierte ich an der Fachhochschule für Sozialpädagogik in Düsseldorf. Studienbegleitend arbeitete ich in verschiedenen sozialpädagogischen Praxisfeldern. Nach dem Erwerb der staatlichen Anerkennung begann ich 1976 das Studium in Erziehungswissenschaft (Hauptfach), Soziologie und Psychologie (Nebenfächer) an der Universität Düsseldorf.